SOME TIME-DEPENDENT QUEUEING PROBLEMS WITH BATCH ARRIVALS AND DEPARTURES

BY

D.R. McNEIL, B.Sc. (Hons),

submitted in fulfilment of the requirements
for the Degree of
Master of Science

UNIVERSITY OF TASMANIA
HOBART

November, 1964

Except as stated herein, this thesis contains no material which has been accepted for the award of any other degree or diploma in any University, and, to the best of the author's knowledge and belief, contains no copy or paraphrase of material previously published or written by another person, except when due reference is made in the text of the thesis.

Signed: D McNeil

(D.R. McNeil)

PREFACE

The research work embodied in this thesis was done during the years 1963-1964 at the Department of Mathematics, the University of Tasmania. The following parts of the thesis are original: Chapter 2, section C, Chapter 3, section D, part 4, Chapter 4, section C. I have submitted for publication some of these results, and it is hoped that they will be published in the near future.

I am indebted to the following: Professor L.S. Goddard who supervised my work in 1963; Professor D. Elliott who helped me considerably with the layout of the thesis; and in particular to Mr. A.M. Hasofer who untiringly assisted and encouraged me during the preparation.

CONTENTS

Page

SUMMARY

In this thesis a queue with infinitely many states, compound Poisson arrivals, bulk service, and batch departures is investigated. It is shown that the queue-length probabilities satisfy an infinite system of differential-difference equations (the "birth and death equations"), which are solved in various special cases. A closely allied system of equations is found for the probability distribution function (which may be defective) of the server's busy period.

In chapter 1 the queue is specified, in terms of the arrival process, the service discipline, and the departure process. The birth and death equations are then derived in their most general form.

In chapter 2 queues with batch arrivals and departures are investigated. This particular case arises by taking the general queue and making the arrival and departure processes independent of the state of the queue. It is here that most of the original work appears, as the finite-time behaviour of queues with batch departures does not seem to have been studied in the literature.

Chapter 3 embodies an exposition of two papers, each by Karlin and McGregor, which the author has studied in detail. In this case the arrival and departure processes depend upon the state of the queue.

In the final chapter the above special cases are considered when $t \to \infty$. In this case, expressions for the probability distribution of a customer's waiting time are also found.

INTRODUCTION

The mathematical theory of queues has received a great deal of attention in recent years. In 1957 a bibliography by Miss A. Doig [14] lists about seven hundred papers on queueing theory, and since then almost as many again have been published. The notion of a queue arises whenever we have individuals arriving, or joining a system in some way, being "served" or simply remaining in the system for a time, and then leaving the system in some manner. The problem usually is: given the manner in which the individuals arrive, are served, and depart, and the initial number in the system, how many will there be in the system after a time t, say, has elapsed? To be able to answer such a question is obviously of practical importance, for all sorts of phenomena may be included in the above broad definition of a queue - a telephone switchboard, customers in a super-market, aeroplanes landing at an aerodrome, bacteria reproducing, vehicles at a traffic interchange, and water in a dam, to quote a few examples. And many modes of behaviour could be included in the processes of arrival, service discipline, and departure for a queue. Individuals may arrive singly or in batches, at regular or random intervals; there may be one, several, or even an infinity of servers; the individuals may be served separately or in bulk, some may become impatient and leave before they are served, and so on. Again only a finite waiting line may be allowed. When one considers all these ramifications, it is not surprising that so much has been written about the theory of queues.

Possibly the simplest example of a queue arises when "customers" arrive at regular intervals and are served separately by a single server, the service times being of constant length. In this case the mathematical model is completely deterministic, and it is a simple matter to work out exactly how many customers there will be in the queue at any given instant. If, in this example, customers were arriving at random instead of regularly, we would have a non-deterministic (stochastic) model, and it would not be possible to determine exactly the number in the queue at a given instant. However it is possible in this case to determine the probability that there will be n customers, say, in the queue at a certain instant (the queue-length probabilities). Surprisingly, for this example it is much more difficult to work out the queue-length probabilities than for the case when the arrival and departure processes are random. In fact, it seems to be a general axiom that mixtures of deterministic and stochastic processes are much more difficult to handle mathematically than purely stochastic processes.

While the subject of queueing theory would seem to be of great practical importance, possibly its main impetus has come not from its usefulness to economists and operations research workers, but because of the wide variety of interesting mathematical techniques it affords. The mathematical model for a queueing situation usually gives rise to either an integral equation or a system of differential-difference equations, but the method of solving may involve the theory of matrices, orthogonal polynomials, Markov chains, Laplace or Stieltjes transforms, complex variables, combinational analysis, alternants and Bessel functions.

Many queueing situations do not easily submit to mathematical analysis, and as a result much of what has been written about the subject is lost in a maze of complicated formula and rather vague theorems. In this thesis we shall be more interested in arriving at a unified theory, giving rise to various known results as special cases, than in the practical use of the results. The particular type of queue investigated is one which lends itself to such a treatment, and may be described as follows:

We begin with an arrival process which consists of a random stream of individuals arriving at some point. By "random" we mean that the probability of a new arrival in an infinitesimally small interval depends only on the length of the interval, and is independent of what has happened before. We generalise this arrival process in two ways:

(a) we allow for individuals to arrive in batches of variable size, by saying that the previously considered individuals were actually groups of individuals;

(b) we allow for the frequency of the incoming stream to depend upon the number of individuals already in the system. Thus we are able to discuss queues in which prospective customers are either deterred or encouraged by the sight of a large queue.

We specify what happens to the individuals after they have arrived by saying that they are served, either separately or in groups of a certain size by a number of servers. The only restriction we make is that the service periods be of a special type of random variable, namely negative exponential. This will automatically define the departure process.

The problem we shall attempt to solve for this particular queueing situation is to determine the probability distribution of two variables, namely, the number of individuals in the system after a given time, and the period during which the server (or servers) is continuously busy. For each of these distributions we shall obtain a set of differential-difference equations. These equations are solved for the above-mentioned special arrival processes (a) and (b) by quite different methods. We shall find that the nature of the solution is governed in the first case by the behaviour of the zeros of a certain polynomial, while the second case involves a study of the properties of certain orthogonal polynomials.

CHAPTER 1

A DESCRIPTION OF THE QUEUE

A. THE ARRIVAL PROCESS

One may formulate a queueing problem by specifying the following three processes:

(a) the manner in which customers arrive,

(b) the discipline under which the customers are served, and

(c) the manner in which the customers depart.

For example, customers may arrive in a purely random manner, be served by a single server in the order "first come first served", and depart from the system a fixed time after commencing service.

The same queueing problem may be formulated in different ways. Suppose we wish to specify the arrival process. If customers arrive at the queue separately, it could be specified by the probability distribution function of the inter-arrival times, that is, by the function

(1.1) $$F_\tau(x) = Prob(\tau \le x)$$

where τ is the time between successive arrivals. However if customers sometimes arrive simultaneously, we would also need to know the probability distribution function of another random variable, r , which is the number of customers in an arriving "batch". Since r can only be a positive integer, it may be specified by the function

(1.2) $$p_r(n) = Prob(r = n)$$

which is often called the probability mass function.

On the other hand, the arrival process could be specified by the probability mass function of the number of customers who have arrived up to time t, that is, by the function

$$p_{A(t)}(n) = Prob(A(t) = n) \tag{1.3}$$

where $A(t)$ is the number of customers who arrive in the interval $(0, t]$. A possible realisation of the kind of arrival process we have been considering is shown in figure 1.

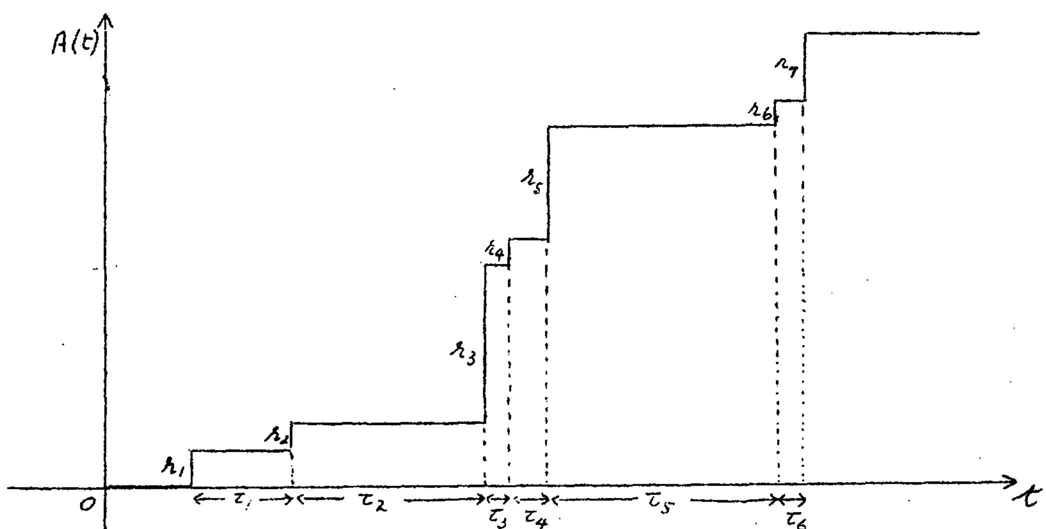


Figure 1.

A third way of specifying this arrival process is by the probability mass function of number of customers who arrive in an interval $(t, t+\delta t]$. For instance we could define

$$(1.4) \qquad Prob\,(A(t+\delta t)-A(t)=r) = \lambda_r \delta t + o(\delta t),$$

$$(1.5) \quad \text{and} \quad Prob\,(A(t+\delta t)-A(t)=0) = 1-\lambda_0\delta t + o(\delta t),$$

where $\{\lambda_r\}$ is a sequence of non-negative terms and $\lambda_0 = \sum_{r=1}^{\infty} \lambda_r$ is finite and strictly positive. Suitable choice of the functions $o(\delta t)$ will then ensure that all the probabilities add up to unity. Suppose we make the further assumption that the arrival process has independent increments, that is, for all choice of indices $t_0 < t_1 < \cdots < t_n$ the n random variables $A(t_1)-A(t_0), \cdots, A(t_n)-A(t_{n-1})$ are independent. This condition is said to make the arrival process additive. Let us show that the functions $F_\tau(x)$, $p_r(n)$ and $p_{A(t)}(n)$, defined in equations (1.1), (1.2) and (1.3) respectively, are expressible in terms of the sequence $\{\lambda_r\}$, defined in equations (1.4) and (1.5).

Suppose there is an arrival at time t_1. Let τ be the time which elapses before the next arrival. Then no customer arrives in the interval $(t_1, t_1+x]$ if and only if $\tau > x$. It follows that

$$(1.6) \qquad Prob\,(\tau > x) = Prob\,(A(t_1+x)-A(t_1)=0).$$

Subdividing the interval $(t_1, t_1+x]$ into $\frac{x}{h}$ subintervals each of length h, we have from equation (1.6)

$$Prob\,(\tau > x) = \sum_{n=1}^{x/h} Prob\,(A(t_1+nh)-A(t_1+\overline{n-1}h)=0).$$

Using equations (1.5) and (1.1) we get

$$F_\tau(x) = 1 - (1 - \lambda_0 h + o(h))^{x/h}.$$

Upon letting $h \to 0$, we finally obtain

(1.7) $$F_\tau(x) = 1 - e^{-\lambda_0 x}.$$

Equation (1.7) specifies the probability distribution function of the inter-arrival times in terms of $\{\lambda_r\}$. The distribution is known as the negative exponential distribution, with mean value $\frac{1}{\lambda_0}$. Obviously the third method of specifying the arrival process, that is, by equations (1.4) and (1.5), places quite a restriction on $F_\tau(x)$. The conditional probability that r customers arrive in the interval $(t, t+\delta t]$, given that at least one customer arrives in the interval is, from equations (1.4) and (1.5)

$$\frac{Prob\,(A(t+\delta t)-A(t) = r\,)}{Prob\,(A(t+\delta t)-A(t) > 0\,)} = \frac{\lambda_r \delta t + o(\delta t)}{\lambda_0 \delta t + o(\delta t)}.$$

This probability tends to $\frac{\lambda_r}{\lambda_0}$ as $t \to 0$, and so the conditional probability that r customers arrive simultaneously at a given instant at which there is an arrival is $\frac{\lambda_r}{\lambda_0}$. It follows from equation (1.2) that r, the size of the arriving batches, has probability mass function

(1.8) $$p_r(n) = \frac{\lambda_n}{\lambda_0}.$$

Finally let us deduce the function $p_{A(t)}(n)$. We will find $p_{A(t)}(n)$ by considering various ways in which the situation "$A(t+\delta t) = n$"

may arise. These are the possibilities:

(i) $A(t) = n$; no customers arrive in the interval $(t, t+\delta t]$;

(ii) $A(t) = n-r$; r customers arrive in the interval $(t, t+\delta t]$, and $r = 1, 2, \cdots n$.

Now using equations (1.4) and (1.5), the probabilities of these events are

$$p_{A(t)}(n)\{1 - \lambda_0 \delta t + o(\delta t)\},$$

$$\text{and} \quad \sum_{r=1}^{n} p_{A(t)}(n-r)\{\lambda_r \delta t + o(\delta t)\},$$

respectively. Hence

$$(1.9) \quad p_{A(t+\delta t)}(n) = p_{A(t)}(n)\{1 - \lambda_0 \delta t + o(\delta t)\} + \sum_{r=1}^{n} p_{A(t)}(n-r)\{\lambda_r \delta t + o(\delta t)\}.$$

Now subtracting $p_{A(t)}(n)$ from each side of equation (1.9) and dividing through by δt, we obtain

$$(1.10) \quad \frac{p_{A(t+\delta t)}(n) - p_{A(t)}(n)}{\delta t} = -\lambda_0 p_{A(t)}(n) + \sum_{r=1}^{n} \lambda_r p_{A(t)}(n-r) + o(\delta t).$$

Now let $\delta t \to 0$. Equation (1.10) becomes

$$(1.11) \quad \frac{\partial}{\partial t} p_{A(t)}(n) = -\lambda_0 p_{A(t)}(n) + \sum_{r=1}^{n} \lambda_r p_{A(t)}(n-r).$$

We define a generating function $\Pi(z,t)$ by

$$(1.12) \quad \Pi(z,t) = \sum_{n=0}^{\infty} z^n p_{A(t)}(n).$$

Multiplying equation (1.11) by z^n and summing over all n we find

(1.13) $$\frac{\partial}{\partial t} \Pi(z,t) = \left(-\lambda_0 + \sum_{r=1}^{\infty} \lambda_r z^r\right) \Pi(z,t).$$

Using the condition that $A(0) = 0$, the solution to the partial differential equation (1.13) is

(1.14) $$\Pi(z,t) = \exp\left(-\lambda_0 + \sum_{r=1}^{\infty} \lambda_r z^r\right) t.$$

Equations (1.12) and (1.14) completely specify the probability mass function $p_{A(t)}(n)$ in terms of the sequence $\{\lambda_r\}$. The arrival process whose generating function is of the form given by equation (1.14) is known as a <u>generalised Poisson process</u>, (see, for example, Parzen [57], p126).

The kind of arrival process we have been describing is quite independent of the other two processes which specify the queueing problem. Since this is not always true we must generalise our arrival process in some way. For the particular kind of arrival process which is defined by equations (1.4) and (1.5), we could do this simply by allowing the constants λ_r to depend on n, the total number of customers in the system, both waiting and being served. We would then have

(1.15) $$\mathrm{Prob}\,(A(t+\delta t) - A(t) = r) = \lambda_{rn}\delta t + o(\delta t),$$

(1.16) $$\text{and } \mathrm{Prob}\,(A(t+\delta t) - A(t) = 0) = 1 - \lambda_{0n}\delta t + o(\delta t).$$

It is possible to generalise further the arrival process by making

λ_{rn} a function of t, in which case we have a non-stationary arrival process. Further mention of this will be made in section D of this chapter (see p27). However in this thesis the most general arrival process which we will consider is specified by equations (1.15) and (1.16).

B. THE SERVICE DISCIPLINE

It will be assumed that the customers form a waiting line, in the order in which they arrive, behind a "counter". In the case of two or more customers arriving simultaneously, some criterion is used to order them in the waiting line. It will also be assumed that the customers remain in the waiting line until it is their turn to be served.

In order that the service discipline be determined, we must specify the following:

(a) The number of customers which can be served simultaneously;

(b) The manner in which customers are served, that is, individually or in batches;

(c) What happens when a server finds there are fewer customers in the waiting line than he has chosen to serve.

We will suppose that, in addition to the customers in the waiting line, the system contains a fixed number, m, of customers receiving service. There are two quite different queueing situations which may arise. The first is the queue with m servers, who serve customers individually, while the second is the queue with a single server who may serve up to m customers simultaneously. One could consider the more general situation where we have r servers, say, each able to serve m customers simultaneously. However we will restrict ourselves to the two cases mentioned above. In the first case it will be assumed that the servers are identical, and act independently of each other.

In order to completely specify the service discipline we must state what happens in each case when there are fewer than m customers in the whole system. For the case of m servers, it simply means that a certain number of the servers will be idle. The case of the single server is more complicated, and will be stated as follows: As soon as the server becomes idle he takes sufficient customers from the waiting line to make m in all when service commences. If, upon becoming idle, the server finds insufficient customers in the waiting line to fulfil his capacity he takes one of two alternatives:

(i) the server takes all the customers in the waiting line and commences service; if more customers arrive during this service time they join those being served, but the number being served never exceeds m;

(ii) the server remains idle until there are sufficient customers in the waiting line to make a total of at least m in the system.

The probability of the first alternative is a fixed number q. It follows that the second alternative has probability $1-q$.

While the above definitions are obviously restrictive, most of the queueing situations which have been discussed in the literature can be specified in this way. Some notable exceptions are the queue with service order "last come, first served", which has been investigated by Wishart [67], the queue with "balking" or "reneging" which Haight [25] considered, and the queue with two non-symmetric channels which was also investigated by Haight [26].

C. THE DEPARTURE PROCESS

Unlike the arrival process and the service discipline, the departure process cannot be made independent of the number of customers in the system. For example, if there are no customers in the system, then a departure is impossible. In the case of the single server, one could define the departure process by specifying

(a) the probability distribution function of the service time, ν, of the server, namely

$$F_\nu(x) = Prob(\nu \leq x), \quad and \tag{1.18}$$

(b) the probability mass function of the number, s, of customers who depart simultaneously at the completion of a service time,

$$p_s(n) = Prob(s = n). \tag{1.19}$$

In the case of m servers, it would be sufficient to specify the probability distribution function of the service time of each server.

If the service time, ν , always follows a negative exponential distribution, with constant mean value $\frac{1}{\mu_0}$ say, where $\mu_0 > 0$, then we have

$$\text{(1.20)} \qquad Prob\,(\nu \leq x) = 1 - e^{-\mu_0 x}.$$

Suppose a particular server is busy at time t . Then the probability that the server completes service in the interval $(t, t+\delta t]$ is

$$Prob\,(\nu \leq t+\delta t / \nu > t) = \frac{Prob\,(t < \nu \leq t+\delta t)}{Prob\,(\nu > t)}.$$

Now using equation (1.20),

$$\begin{aligned} Prob\,(t < \nu \leq t+\delta t) &= 1 - e^{-\mu_0(t+\delta t)} - (1 - e^{-\mu_0 t}), \\ &= e^{-\mu_0 t}(1 - e^{-\mu_0 \delta t}). \end{aligned}$$

Hence

$$\begin{aligned} Prob\,(\nu \leq t+\delta t / \nu > t) &= \frac{e^{-\mu_0 t}(1 - e^{-\mu_0 \delta t})}{1 - e^{-\mu_0 t}}, \\ &= \mu_0 \delta t + o(\delta t). \end{aligned}$$

It follows that the probability of the server completing service more than once in the interval $(t, t+\delta t]$ is $o(\delta t)$.

We have shown that if a server is busy at time t, the probability of a departure from that server in the interval $(t, t+\delta t]$ is $\mu_0 \delta t + o(\delta t)$. Whether or not a particular server is busy at time t, depends on the state of the system.

Now in the case of the queue with m identical servers, all the servers are busy if the number in the system is not less than m, otherwise the number of busy servers is equal to the number of customers in the system. Suppose there are n customers in the system at time t. Then since the servers are independent, we have

$$(1.21) \quad \text{Prob}\,(\text{a departure occurs in the interval } (t, t+\delta t]) = \begin{cases} m\mu_0 \delta t + o(\delta t); & n \geq m \\ n\mu_0 \delta t + o(\delta t); & n < m \end{cases}$$

We have shown that the probability that a particular server completes service twice in an interval $(t, t+\delta t]$ is $o(\delta t)$. Then since each server serves one customer at a time,

$$(1.22) \quad \text{Prob}\,(\text{more than one departure occurs in the interval } (t, t+\delta t]) = o(\delta t).$$

For the queue with a single server the server may serve up to m customers simultaneously, and so there may be up to m simultaneous departures at a given instant. Suppose the server is serving m customers. This is true if and only if the number in the system is not less than m. Then if s is the number of customers who depart simultaneously at end of the service time, we define

$$\text{(1.23)} \qquad Prob\ (s = k) = \frac{\mu_k}{\mu_0},$$

where $\mu_k = 0$ if $k > m$, and $\sum_{k=1}^{m} \mu_k = \mu_0$.

If, on the other hand, the number of customers in the system, n, is less than m, and the server is busy, we define

$$\text{(1.24)} \qquad Prob\ (s = k) = \begin{cases} \frac{\mu_k}{\mu_0} & \text{if } k < n \\ \sum_{j=n}^{m} \frac{\mu_j}{\mu_0} & \text{if } k = n \end{cases}$$

Now the probability of at least one departure in the interval $(t, t+\delta t]$, given the server is busy, is $\mu_0 \delta t + o(\delta t)$. Since the probability that the server is busy is q if the number in the system, n, is less than m, and 1 if $n \geq m$, we obtain, using equations (1.23) and (1.24),

$$\text{(1.25)} \qquad Prob(s \text{ customers depart in the interval } (t, t+\delta t]) = \begin{cases} \mu_s \delta t + o(\delta t); & n \geq m \\ q \sum_{j=n}^{m} \mu_j \delta t + o(\delta t); & s = n < m \\ q \mu_s \delta t + o(\delta t); & s < n < m \\ 0; & n = 0. \end{cases}$$

As in the case for the arrival process, we may wish to allow the service times themselves to depend upon the state of the system. This could be done by allowing the constants $\mu_s, s = 1, 2, \cdots, m$, to depend on n. We would then replace equation (1.25) by

$$\text{(1.26)}\quad \mathrm{Prob}\,(s \text{ customers depart in the interval } (t, t+\delta t]) = \begin{cases} \mu_{sn}\delta t + o(\delta t) \;; & n \geq m \\ q\sum_{j=n}^{m}\mu_{jn}\,\delta t + o(\delta t); & s = n < m \\ q\mu_{sn}\delta t + o(\delta t); & s < n < m \\ 0 \;; & n = 0. \end{cases}$$

Equation (1.21) may be conveniently considered as a special case of equation (1.26) in which

$$\mu_{sn} = \begin{cases} \delta_{s1}\, m\mu_0 & \text{if } n \geq m; \\ \delta_{s1}\, n\mu_0 & \text{if } n < m. \end{cases}$$

It follows from equation (1.26) that

$$\text{(1.27)}\quad \mathrm{Prob}\,(\text{no customers depart in the interval } (t, t+\delta t]) = \begin{cases} 1 - \mu_{0n}\delta t + o(\delta t); & n \geq m \\ 1 - q\mu_{0n}\delta t + o(\delta t); & 1 \leq n < m \\ 1 \;; & n = 0. \end{cases}$$

D. THE BIRTH AND DEATH EQUATIONS

A basic queueing problem is to find the probability that there is a certain number, say n in the system at time t. We shall use the notation

$$\mathrm{Prob}\,(N(t) = n \,/\, N(0) = i) = P_{in}(t)$$

to describe this probability, which is a conditional probability, since it depends on the number, i, of customers in the system at $t = 0$. $N(t)$ is the number in the system, both waiting and receiving service, at time t.

Using the assumptions (1.15) and (1.16) for arrivals, and (1.26) and (1.27) for departures, it is possible to write down a set of differential-difference equations involving $P_{in}(t)$. Let us consider the situation

$$N(t+\delta t) = n.$$

It can arise in the following mutually exclusive and exhaustive ways:

(i) $N(t) = n$; no arrivals or departures in the interval $(t, t+\delta t]$;

(ii) $N(t) = n-r$; r arrivals and no departures in the interval $(t, t+\delta t], r=1,2,\cdots,n$;

(iii) $N(t) = n+s$; no arrivals and s departures in the interval $(t, t+\delta t], s=1,2,\cdots,m$;

(iv) $N(t) = n-r+s$; r arrivals and s departures in the interval $(t, t+\delta t], s=1,2,\cdots,m; r=1,2,\cdots,s+n$.

Using equations (1.15), (1.16), (1.26) and (1.27), the probabilities of these events are, for $n \geq 1$,

$$\text{(i)} \quad \begin{array}{l} P_{in}(t)\{1-\lambda_{on}\delta t+o(\delta t)\}\{1-\mu_{om}\delta t+o(\delta t)\} \text{ if } n \geq m, \\ P_{in}(t)\{1-\lambda_{on}\delta t+o(\delta t)\}\{1-q\mu_{on}\delta t+o(\delta t)\} \text{ if } n < m; \end{array}$$

(ii) $\sum_{r=1}^{n} P_{i,n-r}(t)\left\{\lambda_{r,n-r}\delta t + o(\delta t)\right\} + o(\delta t)$;

(iii) $\sum_{s=1}^{m} P_{i,n+s}(t)\left\{\mu_{sm}\delta t + o(\delta t)\right\} + o(\delta t)$ if $n \geq m$,

$\sum_{s=1}^{m-n-1} P_{i,n+s}(t)\left\{q\mu_{s,n+s}\delta t + o(\delta t)\right\} + \sum_{s=m-n}^{m} P_{i,n+s}(t)\left\{\mu_{sm}\delta t + o(\delta t)\right\} + o(\delta t)$ if $n < m$;

(iv) and $\sum_{s=1}^{m}\sum_{r=1}^{s+n} P_{i,n-r+s}(t) \cdot o(\delta t)$; respectively .

Now the probability of the event "$N(t+\delta t) = n$" is $P_{in}(t+\delta t)$. So adding the probabilities we get:

$$P_{in}(t+\delta t) = P_{in}(t) - (\lambda_{on} + \mu_{on})\delta t P_{in}(t) + \sum_{r=1}^{n}\lambda_{r,n-r}\delta t P_{i,n-r}(t) + \sum_{s=1}^{m}\mu_{sm}\delta t P_{i,n+s}(t) + o(\delta t); \quad n \geq m, \tag{1.28}$$

$$P_{in}(t+\delta t) = P_{in}(t) - (\lambda_{on} + q\mu_{on})\delta t P_{in}(t) + \sum_{r=1}^{n}\lambda_{r,n-r}\delta t P_{i,n-r}(t) + q\sum_{s=1}^{m-n-1}\mu_{s,n+s}\delta t P_{i,n+s}(t) + \sum_{s=m-n}^{m}\mu_{sm}\delta t P_{i,n+s}(t) + o(\delta t); \quad n < m. \tag{1.29}$$

When $n = 0$, event (ii) is impossible. Using equation (1.26) we get

$$P_{io}(t+\delta t) = P_{io}(t) - \lambda_{oo}\delta t P_{io}(t) + q\sum_{s=1}^{m-1}\sum_{j=s}^{m}\mu_{js}\delta t P_{is}(t) + \mu_{mm} P_{im}(t) + o(\delta t). \tag{1.30}$$

Now subtracting $P_{in}(t)$ from each side of equations (1.28), (1.29) and (1.30), dividing through by δt , and letting $\delta t \to 0$, we obtain the differential-difference equations

$$(1.31)\quad \frac{\partial}{\partial t} P_{in}(t) = \sum_{r=1}^{n} \lambda_{r,n-r} P_{i,n-r}(t) - (\lambda_{on}+\mu_{on}) P_{in}(t) + \sum_{s=1}^{m} \mu_{sm} P_{i,n+s}(t);\; n \geq m,$$

$$(1.32)\quad \frac{\partial}{\partial t} P_{in}(t) = \sum_{r=1}^{n} \lambda_{r,n-r} P_{i,n-r}(t) - (\lambda_{on}+q\mu_{on}) P_{in}(t) + q\sum_{s=1}^{m-n-1} \mu_{s,n+s} P_{i,n+s}(t) + \sum_{s=m-n}^{m} \mu_{sm} P_{i,n+s}(t);\; 0 < n < m,$$

$$(1.33)\ \text{and}\quad \frac{\partial}{\partial t} P_{i0}(t) = -\lambda_{00} P_{i0}(t) + q\sum_{s=1}^{m-1}\sum_{j=s}^{m} \mu_{js} P_{is}(t) + \mu_{mm} P_{im}(t).$$

These equations are a special case of what Bharucha-Reid [5,p60] calls the "Feller integrodifferential equations", since they were first obtained by Feller [20].

The system of differential-difference equations (1.31) and (1.33) together with the initial condition

$$(1.34)\qquad P_{in}(0) = \delta_{in}$$

has been investigated, in special cases, by many authors. For example if we put $\lambda_{rn} = \delta_{r1}\lambda_n$, and $\mu_{sn} = \delta_{s1}\mu_n$, we have the system of equations which has been investigated in great detail by Ledermann and Reuter [48] and Karlin and McGregor [31]. Less has been written about the case when $\lambda_{rn} = \lambda_r$ and $\mu_{sn} = \mu_s$, although Luchak [50] considered the particular case when $\mu_s = \delta_{s1}\mu$, and Foster [22], Bailey [3], Wishart [68], Downton [15] and others investigated the behaviour of the $P_{in}(t)$ as $t \to \infty$, for more general inter-arrival and service time distributions.

The equations which Karlin and McGregor investigated are known as the "birth and death" equations, and describe systems other than queues. In fact the equations first arose in the study of populations, whence the terms "birth" and "death". The more general equations have not been written down as such in the literature to date. In the remainder of this thesis equations (1.31), (1.32) and (1.33) will be called the birth and death equations.

The kind of queue which we have been describing is a Markov process. A Markov process has the property that the probability law of the future development of the process, once it is in a given state, depends only on the state and not on how the process arrived in that state. Since equations (1.31),(1.32) and (1.33) only involve the functions $P_{in}(t)$ at the instant t, and the integer i, then the state of the queue for all t is completely specified by i, and consequently this queue is a Markov process.

There are two kinds of queueing problems, in addition to the ones already described, which may be described by a system of differential-difference equations similar to equations (1.31), (1.32) and (1.33).

(a) The first is the queue in which the respective probabilities of an arrival and a departure in the interval $(t, t+\delta t]$ are functions of t. This is called the queue with non-stationary parameters, and has been

studied by Clarke [8], and Keilson [40], in the special case when there is a single server and $m = 1$. Using different methods, these authors showed that the $P_{in}(t)$ satisfy a Volterra-type integral equation. However no example has been worked out and expressed in terms of known functions, and it seems that the mathematics is very intractable for this particular queue. Indeed the only non-trivial example of a queue with non-stationary parameters for which the $P_{in}(t)$ can be found explicitly is a queue with an infinite number of servers (see Von Sydow [64]).
However the mathematics in this case is not markedly different in principle from the case for the queue with stationary parameters, which we will consider in chapter 3, so we will not consider it.

(b) The second type of queue representable by a system of differential-difference equations is that in which only a finite waiting line is allowed. Since the system consequently has only a finite number of states, there are only a finite number of difference equations. Again the mathematics is not very tractable. However Ledermann and Reuter [48] have considered this problem for the case $\lambda_{rn} = \delta_{r1}\lambda_n$ and $\mu_{sn} = \delta_{s1}\mu_n$, and shown that the solution may be expressed in terms of the eigenvalues of a matrix. In the case when the capacity of the system is small Saaty [59,ch5] gives a good account of the matrix methods which are involved in obtaining the solution. By taking Laplace transforms in the difference equations, Moran [54] expressed the Laplace transforms of the $P_{in}(t)$ as quotients of certain polynominals. In the case when $\lambda_n = \lambda$ and $\mu_n = \mu$ for all n, expressions for the $P_{in}(t)$ in terms of infinite series of modified Bessel functions were obtained by Takacs [62,p13] and also by Heathcote and Moyal [27], while Morse [55] expressed the result in terms of trigonometric functions.

In this thesis neither of the above two types of problem will be investigated, and we shall confine our attention to queues which can be described by the birth and death equations (1.31), (1.32) and (1.33). We will distinguish between two special types, which shall be the subject matter of chapters 2 and 3 respectively and in each of which the parameters λ_{rn} and μ_{sn} take a special form:

(i) The first is the queue with a single server in which

$\lambda_{rn} = \lambda_r$ and $\mu_{sn} = \mu_s$, for all n.

Since the parameters λ_{rn} and μ_{sn}, which characterise the arrival and departure processes respectively, are in this case independent of the state of the system, n, we shall call this queue the queue with <u>state-independent</u> parameters.

(ii) The second type is the queue in which $\lambda_{rn} = \delta_{r1}\lambda_n$ and

$\mu_{sn} = \delta_{s1}\mu_n$.

Since $\lambda_{rn} = \mu_{sn} = 0$ unless $r = s = 1$, this is the queue in which no two or more arrivals or departures can occur simultaneously. Accordingly we shall call this queue the queue with single arrivals and departures, or the queue with simple parameters.

A very important queue is the one with both state-independent and simple parameters. This is the queue with a single server and negative exponential distributions of inter-arrival and service times. Kendall [42] called it the queue $M/M/1$, where the first letter labels the arrival process, which is random (that is, the inter-arrival times are negative exponential variables), the second letter the departure process, and the number 1 refers to the number of servers. Ledermann and Reuter [48], Bailey [2], and Champernowne [7], and Conolly [9], using different methods, obtained an expression for $P_{in}(t)$ as an infinite series of modified Bessel functions, while Karlin and McGregor [33] and Morse [55] each obtained a different integral representation for the result. Throughout this thesis we will obtain these expressions as special cases of queues of types (i) and (ii).

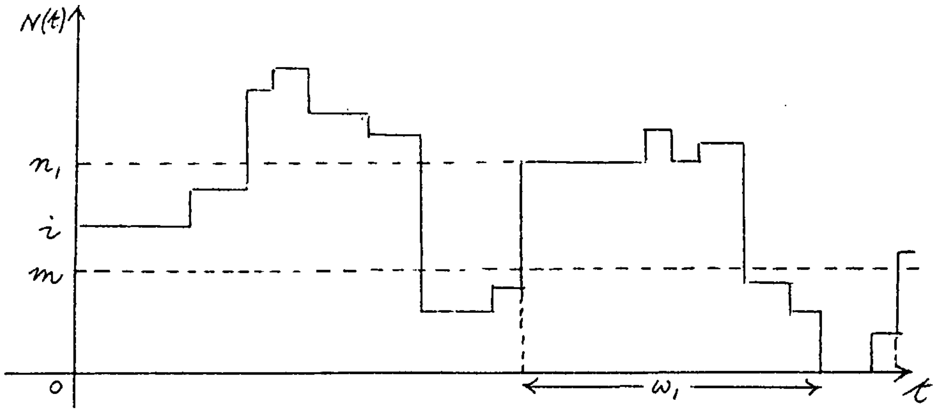
E. THE BUSY PERIODS

We have seen that there are certain times when one or more of the servers in the queue are idle. In the single server queue, for example, whenever the number in the system drops below m, there is a probability $1-q$ that the server remains idle until there are at least m customers in the system. For this queue, we will define a period ω, during which the server is continuously busy, as a <u>busy period</u>.

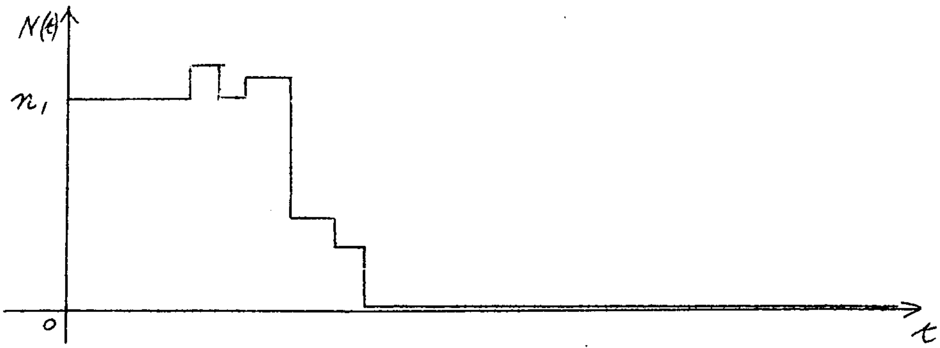
Bailey [2] obtained a simple expression for the probability density function (the derivative of the probability distribution function) of ω, for the queue $M/M/1$. His method was later applied to a queue with non-simple arrival parameters (arrivals in batches) by Luchak [50]. We will further generalise Bailey's method to obtain the distribution function of ω for the single server queue we have been considering.

We consider the modified process which commences at the start of the first busy period, and which ceases for all time as soon as the server becomes idle. The relationship between this process and the original queueing process can be visualised by considering one particular realisation of the graph of $N(t)$ against t for the two processes, as in figure 2.

The modified process commences when the queue is in some state $i_1 \geq m$, and ceases when the queue is in a state $j_1 < m$. In the case we have drawn $j_1 = 0$, but this is not true in general, because there is a probability $1-q$ that the server remains idle when the number in the system is less than m.



A REALISATION OF THE ORIGINAL QUEUEING PROCESS



A REALISATION OF THE MODIFIED PROCESS

Figure 2.

The duration of the first busy period, ω_1 , is less than or equal to t if, and only if, for the modified process the server is idle at time t . Hence, if $\hat{P}_{i,n}(t)$ denotes the probability that the modified process is in the state n at time t , given it commenced in the state i_1 ,we have

$$(1.35) \quad \text{Prob}\,(\omega_1 \leq t \mid \text{first busy period commenced with } i_1 \text{ in the queue}) = \hat{P}_{i_1,0}(t) + (1-q)\sum_{n=1}^{m-1} \hat{P}_{i_1,n}(t).$$

Now let us suppose that i_1 has the probability mass function

(1.36) $\quad Prob(i_1 = n) = a_n, \quad n = m, m+1, \cdots\cdots$

Then from equation (1.35) we have

(1.37) $\quad Prob(\omega_1 \le t) = \sum_{i=m}^{\infty} a_i \{ \hat{P}_{i0}(t) + (1-q) \sum_{n=1}^{m-1} \hat{P}_{in}(t) \}.$

Now since the parameters λ_{rn} and μ_{sn} describing the arrival and departure process are independent of the time t, all the busy periods have the same distribution. Consequently equation (1.37) gives the probability distribution function of the busy periods, provided the $\hat{P}_{in}(t)$ are known. Let us derive a system of differential equations for the $\hat{P}_{in}(t)$.

For the modified process the assumptions concerning the arrival process are

(1.38) $Prob(r \text{ customers arrive in the interval } (t, t+\delta t]) = \begin{cases} \lambda_{rn}\delta t + o(\delta t); & n \ge m, \\ q\lambda_{rn}\delta t + o(\delta t); & 1 \le n < m, \\ 0 \;; & n = 0, \text{ and} \end{cases}$

(1.39) $Prob(\text{no customers arrive in the interval } (t, t+\delta t]) = \begin{cases} 1 - \lambda_{on}\delta t + o(\delta t); & n \ge m, \\ 1 - q\lambda_{on}\delta t + o(\delta t); & 1 \le n < m, \\ 1 \;; & n = 0. \end{cases}$

The assumptions regarding departures are the same as for the original process, namely, equations (1.26) and (1.27). Using similar arguments to those used to obtain equations (1.31), (1.32) and (1.33), we find

$$(1.40)\quad \frac{\partial}{\partial t}\hat{P}_{in}(t) = \sum_{r=1}^{n-m}\lambda_{r,n-r}\hat{P}_{i,n-r}(t) + q\sum_{r=n-m+1}^{n-1}\lambda_{r,n-r}\hat{P}_{i,n-r}(t) - (\lambda_{0n}+\mu_{0n})\hat{P}_{in}(t) + \sum_{s=1}^{m}\mu_{s,n+s}\hat{P}_{i,n+s}(t),\ n \geq m;$$

$$(1.41)\quad \frac{\partial}{\partial t}\hat{P}_{in}(t) = q\left\{\sum_{r=1}^{n-1}\lambda_{r,n-r}\hat{P}_{i,n-r}(t) - (\lambda_{0n}+\mu_{0n})\hat{P}_{in}(t) + \sum_{s=1}^{m-n-1}\mu_{s,n+s}\hat{P}_{i,n+s}(t)\right\} + \sum_{s=m-n}^{m}\mu_{s,n+s}\hat{P}_{i,n+s}(t),\ 1 \leq n < m;$$

$$(1.42)\quad \frac{\partial}{\partial t}\hat{P}_{i0}(t) = q\sum_{s=1}^{m-1}\sum_{k=s}^{m}\mu_{ks}\hat{P}_{is}(t) + \mu_{mm}\hat{P}_{im}(t);$$

$$(1.43)\quad \text{and}\quad \hat{P}_{in}(0) = \delta_{in}.$$

We specified earlier in this section that the modified process commences at the instant when the first busy period starts. However in obtaining equations (1.40), (1.41) and (1.42) we have only specified that the modified process commences when the number in the system is equal to the number at the start of the busy period. There would seem to be a discrepancy, in that the modified process could commence at any time after the busy period commenced, provided there has been no change in the state of the system. The explanation is that it does not matter, because the times when the state of the system remains unaltered (the service times) are negative exponential variables, ν, which have the property that

$$(1.44)\quad \mathrm{Prob}(\nu > x+y \,/\, \nu > x) = \mathrm{Prob}(\nu > y),$$

and so the busy period has the same probability distribution no matter what the point from which we measure it, provided there is no alteration to the state of the system. The property stated in equation (1.44) is known as the "no-memory" property, and the negative exponential distribution is the only probability distribution which possesses this important property. Because of this property of the service time distribution it also follows that customers who join those being served after service has commenced will have the same service time distribution as customers who are "on time".

There is one important point which we must mention. It is conceivable that the first busy period may be of infinite duration, in which case the function $Prob(\omega \leq x)$ is not a true probability distribution, since

$$Prob(\omega \leq \infty) \neq 1.$$

This could occur, for example, if customers are arriving at a much greater rate than can be coped with by the server. The precise conditions which are necessary in order that the function on the right hand side of equation (1.37) be a probability distribution, will be discussed for each particular case separately, in the next two chapters. A function $F(x)$ which fulfils all the requirements of a probability distribution function except the property

$$\lim_{x \to \infty} F(x) = 1$$

is often called a "dishonest" distribution function.

CHAPTER 2

QUEUES WITH ARRIVALS AND DEPARTURES IN BATCHES

In this chapter we shall investigate various queues which arise when the parameters λ_{rn} and μ_{sn} describing the arrival and departure processes respectively, are independent of n, the state of the queue. These queues are the simplest in which arrivals and departures occur in batches. Very little has been written about the finite-time behaviour of such queues. The main contributions have come from Luchak [50], Gaver [24], and more recently from Keilson [41] and Bhat [6]. However a number of papers have appeared on the asymptotic behaviour of batch queues as $t \to \infty$. We will discuss this aspect more fully in chapter 4.

For each example in this chapter we will obtain expressions for the state probabilities, $P_{in}(t)$, and the probability density function of the busy period,

$$f_w(t) = \frac{\partial}{\partial t} \mathrm{Prob}(w \le t) . \tag{2.1}$$

A. GENERAL THEORY - STATE-INDEPENDENT PARAMETERS

The functions λ_{rn} and μ_{sn} may be replaced by λ_r and μ_s since they do not involve n. The birth and death equations for the state probabilities, (1.31), (1.32) and (1.33), become

$$\frac{\partial}{\partial t} P_{in}(t) = \sum_{r=1}^{n} \lambda_r P_{i,n-r}(t) - (\lambda_0 + \mu_0) P_{in}(t) + \sum_{s=1}^{m} \mu_s P_{i,n+s}(t); \quad n \ge m, \tag{2.2}$$

$$\frac{\partial}{\partial t} P_{in}(t) = \sum_{r=1}^{n} \lambda_r P_{i,n-r}(t) - (\lambda_0 + q\mu_0) P_{in}(t) + q\sum_{s=1}^{m-n-1} \mu_s P_{i,n+s}(t) + \sum_{s=m-n}^{m} \mu_s P_{i,n+s}(t), \quad 0 < n < m, \tag{2.3}$$

(2.4) and
$$\frac{\partial}{\partial t} P_{i0}(t) = -\lambda_0 P_{i0}(t) + q \sum_{s=1}^{m-1} \sum_{j=s}^{m} \mu_j P_{is}(t) + \mu_m P_{im}(t).$$

For the modified process from which we obtain the distribution of the busy periods, equations (1.40), (1.41) and (1.42) become

$$(2.5) \quad \frac{\partial}{\partial t} \hat{P}_{in}(t) = \sum_{r=1}^{n-m} \lambda_r \hat{P}_{i,n-r}(t) + q \sum_{r=n-m+1}^{n-1} \lambda_r \hat{P}_{i,n-r}(t) - (\lambda_0+\mu_0)\hat{P}_{in}(t) + \sum_{s=1}^{m} \mu_s \hat{P}_{i,n+s}(t); \quad n \geq m,$$

$$(2.6) \quad \frac{\partial}{\partial t} \hat{P}_{in}(t) = q\left\{\sum_{r=1}^{n-1} \lambda_r \hat{P}_{i,n-r}(t) - (\lambda_0+\mu_0)\hat{P}_{in}(t) + \sum_{s=1}^{m-n-1} \mu_s \hat{P}_{i,n+s}(t)\right\} + \sum_{s=m-n}^{m} \mu_s \hat{P}_{i,n+s}(t), \quad 1 \leq n < m,$$

$$(2.7) \text{ and } \frac{\partial}{\partial t} \hat{P}_{i0}(t) = q \sum_{s=1}^{m-1} \sum_{k=s}^{m} \mu_k \hat{P}_{is}(t) + \mu_m \hat{P}_{im}(t).$$

Since λ_r and μ_s are independent of n, we will use generating functions to solve equations (2.2) - (2.7). Let us define

$$(2.8) \qquad \Pi(z,t) = \sum_{n=0}^{\infty} z^n P_{in}(t)$$

$$(2.9) \qquad \text{and} \quad \hat{\Pi}(z,t) = \sum_{n=0}^{\infty} z^n \hat{P}_{in}(t).$$

Multiplying equations (2.2) and (2.3) by z^n and summing over all positive integral values of n, adding equation (2.4), and using equation (2.8), we obtain

$$(2.10) \quad \frac{\partial}{\partial t} \Pi(z,t) = \left\{\sum_{r=1}^{\infty} \lambda_r z^r - (\lambda_0+\mu_0) + \sum_{s=1}^{m} \mu_s z^{-s}\right\} \Pi(z,t) - \left(\sum_{s=1}^{m} \mu_s z^{-s} - \mu_0\right) P_{i0}(t)$$
$$- \sum_{n=1}^{m-1} \left\{(1-q)\left(\sum_{s=1}^{m} \mu_s z^{-s} - \mu_0\right) + q \sum_{s=n+1}^{m} \mu_s (z^{-s} - z^{-n})\right\} z^n P_{in}(t).$$

In a similar manner we get for $\hat{\Pi}(z,t)$

$$(2.11) \quad \frac{\partial}{\partial t} \hat{\Pi}(z,t) = \left\{\sum_{r=1}^{\infty} \lambda_r z^r - (\lambda_0+\mu_0) + \sum_{s=1}^{m} \mu_s z^{-s}\right\}\left\{\hat{\Pi}(z,t) - \hat{P}_{i0}(t)\right\}$$
$$- \sum_{n=1}^{m-1} \left\{(1-q)\left(\sum_{r=1}^{\infty} \lambda_r z^r - (\lambda_0+\mu_0) + \sum_{s=1}^{m} \mu_s z^{-s}\right) + q \sum_{s=n+1}^{m} \mu_s (z^{-s} - z^{-n})\right\} z^n \hat{P}_{in}(t).$$

Since (2.10) and (2.11) are partial differential equations whose coefficients are independent of t, we will solve them by taking Laplace transforms. The Laplace transform with respect to time is defined by

$$\varphi^*(p) = \int_0^\infty e^{-pt}\varphi(t)\,dt, \quad Re(p) > 0.$$

Since $0 \leq P_{in}(t) \leq 1$, both $P_{in}^*(p)$ and $\hat{P}_{in}^*(p)$ are convergent provided $Re(p) > 0$. In the appendix (Theorem 1) it is proved that $\Pi^*(z,p)$ and $\hat{\Pi}^*(z,p)$ are both analytic in z for fixed p, and in p for fixed z, provided $|z| \leq 1-\varepsilon$ and $Re(p) \geq \delta$, where $0 < \varepsilon < 1$ and $\delta > 0$. We define

$$f(z) = \sum_{r=1}^{\infty} \lambda_r z^r \quad \text{and} \quad h(z) = \sum_{s=1}^{m} \mu_s z^{-s}. \tag{2.12}$$

Then taking Laplace transforms in equations (2.10) and (2.11), and using the initial conditions

$$\Pi(z,0) = \hat{\Pi}(z,0) = z^i,$$

which follow from equations (1.34) and (1.43), we find

$$\Pi^*(z,p) = \frac{z^i - \{h(z)-\mu_0\}P_{i0}^*(p) - \sum_{n=1}^{m-1}\{(1-q)[h(z)-\mu_0] + q\sum_{s=n+1}^{m}\mu_s(z^{-s}-z^{-n})\}z^n P_{in}^*(p)}{-f(z) + (\lambda_0+\mu_0+p) - h(z)}, \tag{2.13}$$

and

$$\hat{\Pi}^*(z,p) = \frac{z^i - \{f(z)-(\lambda_0+\mu_0)+h(z)\}\hat{P}_{i0}^*(p) - \sum_{n=1}^{m-1}\{(1-q)[f(z)-(\lambda_0+\mu_0)+h(z)] + q\sum_{s=n+1}^{m}\mu_s(z^{-s}-z^{-n})\}z^n \hat{P}_{in}^*(p)}{-f(z) + (\lambda_0+\mu_0+p) - h(z)}. \tag{2.14}$$

In the appendix (Theorem 2) it is proved that the expression

$$-f(z) + (\lambda_0 + \mu_0 + p) - k(z),$$

considered as a function of z, has precisely m zeros in the region $|z| \leq 1-\varepsilon$, $Re(p) \geq \delta$. Let us denote these zeros by $\xi_1, \xi_2, \cdots, \xi_m$. From the analyticity of $\Pi^*(z,p)$ and $\hat{\Pi}^*(z,p)$ in this region it follows that the ξ_r are also zeros of the respective numerators in the right hand sides of equations (2.13) and (2.14). This gives us m equations for $P_{i0}^*(p), P_{i1}^*(p), \cdots, P_{i,m-1}^*(p)$, and m equations for $\hat{P}_{i0}^*(p)$, $\hat{P}_{i1}^*(p), \cdots, \hat{P}_{i,m-1}^*(p)$. The relations for the $P_{ir}^*(p)$ are:

(2.15) $$[k(\xi_r)-\mu_0]P_{i0}^*(p) + \sum_{n=1}^{m-1}\{(1-q)[k(\xi_r)-\mu_0] + q\sum_{s=n+1}^{m}\mu_s(\xi_r^{-s}-\xi_r^{-n})\}\xi_r^{n}P_{in}^*(p) = \xi_r^{i},$$

where $r = 1, 2, \cdots, m$. Using the fact that

$$-f(\xi_r) + (\lambda_0+\mu_0+p) - k(\xi_r) = 0, \quad r = 1, 2, \cdots, m,$$

the equations for the $\hat{P}_{in}^*(p)$ reduce to

(2.16) $$p\hat{P}_{i0}^*(p) + \sum_{n=1}^{m-1}\{(1-q)p + q\sum_{s=n+1}^{m}\mu_s(\xi_r^{-s}-\xi_r^{-n})\}\xi_r^{n}\hat{P}_{in}^*(p) = \xi_r^{i},$$

where $r = 1, 2, \cdots, m$.

The zeros $\xi_1, \xi_2, \cdots, \xi_m$ may not be distinct. Suppose ξ_r is of multiplicity k. Then denoting the left hand side of equation (2.15), for example, by $F(\xi_r)$, the multiplicity of ξ_r will give rise to k equations of the form

$$\frac{d^n}{dz^n}\left[F(z) - z^i\right]_{z=\xi_r} = 0, \quad n = 0, 1, \cdots, k-1.$$

Thus we still have m equations for the $P_{in}^*(p)$ and for the $\hat{P}_{in}^*(p)$, provided, of course, they are independent. Assuming this to be so, we may find the state probabilities $P_{in}(t)$, for $0 \leq n < m$, by solving equations (2.15) and inverting Laplace transforms. It should be possible to find the remaining $P_{in}(t)$ by substitution in the birth and death equations. The probability distribution function of the busy period, $F_\omega(t)$, is found by solving equations (2.16), using equation (1.37). We will work out these expressions for some special examples in the next two sections of this chapter. Before doing this, however, let us find an expression for the mean value of $N(t)$, the number in the system at time t.

Differentiating equation (2.10) with respect to z, using equation (2.8), and putting $z = 1$, we get

$$(2.17) \quad \frac{\partial}{\partial t}\sum_{n=0}^{\infty} n P_{in}(t) = \left\{\sum_{r=1}^{\infty} r\lambda_r - \sum_{s=1}^{m} s\mu_s\right\}\sum_{n=0}^{\infty} P_{in}(t) + \sum_{s=1}^{m} s\mu_s P_{i0}(t) + \sum_{n=1}^{m-1}\left\{(1-q)\sum_{s=1}^{n} s\mu_s - q\sum_{s=n+1}^{m}(n-s)\mu_s\right\}P_{in}(t).$$

The mean value of $N(t)$ is defined by

$$(2.18) \qquad M(t) = \sum_{n=0}^{\infty} n P_{in}(t).$$

Since $\sum_{r=1}^{\infty} r\lambda_r = f'(1)$ and $-\sum_{s=1}^{m} s\mu_s = k'(1)$, from equation (2.12), and $\sum_{n=0}^{\infty} P_{in}(t) = 1$, equation (2.17) reduces to

$$(2.19) \qquad \frac{\partial}{\partial t} M(t) = f'(1) + h'(1) - h'(1)P_{i0}(t) - \sum_{n=1}^{m-1}\Big\{(1-q)h'(1) \cdot q\sum_{s=n+1}^{m}(n-s)\mu_s\Big\}P_{in}(t).$$

This equation may be integrated to obtain the mean value $M(t)$, once the functions $P_{in}(t)$, $n=0,1,\ldots,m-1$ are known. In the special case of single departures, $\mu_s = \delta_{s1}\mu$ and so equation (2.19) reduces to the equation

$$(2.20) \qquad \frac{\partial}{\partial t} M(t) = f'(1) - \mu\big(1 - P_{i0}(t)\big),$$

which Luchak [49] obtained.

B. THE QUEUE WITH BATCH ARRIVALS AND SINGLE DEPARTURES

If we impose the restriction $\mu_s = \delta_{s1}\mu$, we have the queue with batch arrivals and single departures. Luchak [50] obtained exactly the same system of equations but the queue he considered is different. He considered the queue with single arrivals and departures, but with each service time the sum of h identical negative exponential variables, where h is itself a random variable. If, in the queue we are now considering, we consider each batch of arriving customers as a unit, the total service time of the unit is equal to the sum of the service times of the individual customers, these being identically distributed negative exponential variables. Hence Luchak's queue gives rise to the same system of birth and death equations as the queue with batch arrivals. Foster [22] observed this fact when considering the asymptotic behaviour of the $P_{in}(t)$ as $t \to \infty$. Thus the queue with batch arrivals is a particular case of the queue $M/G/1$, where G refers to a general departure process. This queue was first studied by Pollaczek [58], and later by Takacs [62] who observed that while the number of customers in the system is not a Markov process, the time required to complete the service of all customers in the system at time t has the Markov property.

Putting $\mu_s = \delta_{s1}\mu$ in equations (2.15) and (2.16) we find

$$(2.21) \qquad P^*_{i0}(p) = \frac{\xi_1^{i+1}}{\mu(1-\xi_1)},$$

$$(2.22) \quad \text{and} \quad \hat{P}^*_{i0}(p) = \frac{\xi_1^{i}}{p}.$$

Let us consider the busy period distribution first. Since departures occur one at a time, the queue is always empty when the server is not busy, and so the number in the queue at the commencement of a busy period is equal to the number of customers who arrive in a batch; this has distribution

$$a_i = \frac{\lambda_i}{\lambda_0}$$

from equation (1.8). It follows from equation (1.37) that the busy period, ω, has probability distribution function

$$(2.23) \qquad F_\omega(t) = \sum_{i=1}^{\infty} \frac{\lambda_i}{\lambda_0} \hat{P}_{i0}(t).$$

Using equation (2.22), we find that the Laplace transform of $F_\omega(t)$ is

$$(2.24) \qquad F_\omega^*(p) = \sum_{i=1}^{\infty} \frac{\lambda_i \xi_1^i}{\lambda_0 p}.$$

By definition (2.12),

$$\sum_{i=1}^{\infty} \lambda_i \xi_1^i = f(\xi_1) = \lambda_0 + \mu + p - \frac{\mu}{\xi_1},$$

since $-f(z) + (\lambda_0 + \mu + p) - \frac{\mu}{z}$ vanishes when $z = \xi_1$. Thus equation (2.24) may be writtem as

$$(2.25) \qquad F_\omega^*(p) = 1 + \frac{1}{p}\left\{1 - \mu\left(\frac{1}{\xi_1} - 1\right)\right\}.$$

Now using the result that

$$(2.26) \qquad \lim_{t\to\infty} f(t) = \lim_{p\to 0+} p f^*(p),$$

(provided the left-hand side exists - see, for example Doetsch [13, vll,p451]) we have from equation (2.25)

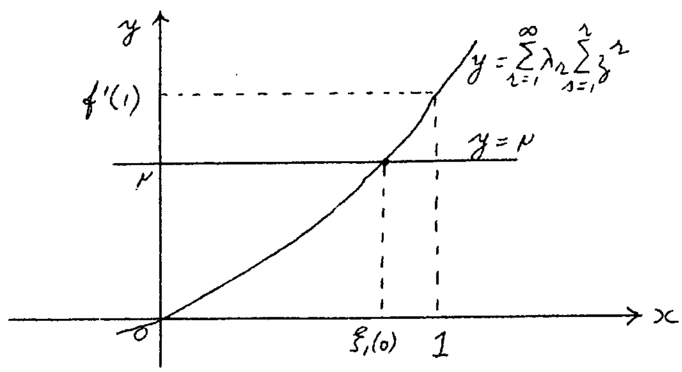
$$(2.27) \qquad \lim_{t\to\infty} F_\omega(t) = 1 - \mu\left(\frac{1}{\xi_1(0)} - 1\right),$$

where $\xi_1(0)$ is the zero of $-f(z) + (\lambda_0 + \mu) - \frac{\mu}{z}$ which has smallest modulus.

<u>THEOREM 2.1</u>: Let $\xi_1(0)$ be the zero of $-f(z) + (\lambda_0 + \mu) - \frac{\mu}{z}$ with the smallest modulus. $\xi_1(0)$ is real and if $f'(1) \le \mu$, $\xi_1(0) = 1$; otherwise $\xi_1(0) < 1$.

<u>PROOF:</u> Let us consider the analytic function

$$\frac{-zf(z)+(\lambda_0+\mu)z-\mu}{1-z} = \sum_{r=1}^{\infty}\lambda_r\sum_{s=1}^{r} z^s - \mu .$$



Since the curves $y=\sum_{r=1}^{\infty}\lambda_r\sum_{s=1}^{r}x^s$ and $y=\mu$ intersect for some x such that $0\le x<1$ if $\sum_{r=1}^{\infty} r\lambda_r>\mu$, the above function has a real zero in the interval $[0,1)$. But we have from the appendix (Theorem 2) that the function $-zf(z)+(\lambda_0+\mu)z-\mu$ has only one zero inside or on the unit circle. Hence if $f'(1)>\mu$, $\xi_1(0)<1$.

If, on the other hand

$$f'(1)\le\mu, \tag{2.28}$$

the inequality

$$\left|\sum_{r=1}^{\infty}\lambda_r\sum_{s=1}^{r}z^s\right| < \mu$$

is satisfied at all points on the contour $|z| = 1-\varepsilon$, and so by Rouche's Theorem (see appendix) the function $\sum_{r=1}^{\infty} \lambda_r \sum_{s=1}^{r} z^r - \mu$ has no zeros in the region $|z| \le 1-\varepsilon$. Consequently $\xi_1(0) = 1$.

We have shown that $\xi_1(0) = 1$ if, and only if, $f'(1) \le \mu$. It follows from equation (2.27) that the function $F_\omega(t)$ is a true probability function if and only if, $f'(1) \le \mu$. If $f'(1) > \mu$, we have, using equation (2.27), that there is a probability

$$\mu \left(\frac{1}{\xi_1(0)} - 1 \right)$$

that a busy period never ends. Now $f'(1)$ is the rate at which customers are arriving. This may be shown as follows: From equation (1.4), the mean number of customers who arrive in an interval of length δt is

$$\sum_{r=1}^{\infty} r \left[\lambda_r \delta t + o(\delta t) \right].$$

Dividing by δt, and using equation (1.5), we find that the mean rate at which customers are joining the system is

$$\sum_{r=1}^{\infty} r \lambda_r = f'(1).$$

Similarly from equation (1.26), putting $\mu_{sn} = \delta_{s1} \mu$, the mean rate at which customers are departing from service is μ. So (2.28) says that the mean arrival rate is less than or equal to the mean rate of departures from service. That this is the condition for the busy periods to have an "honest" distribution is as would be expected. When the inequality (2.28) is not satisfied, the mean arrival rate exceeds the mean rate of departures from service, and again we would expect there to be a non-zero probability that a busy period will never end.

Let us find $f_\omega(t)$, probability density function of the busy period. Since $f_\omega(t) = \frac{d}{dt} F_\omega(t)$, the Laplace transform of $f_\omega(t)$ is

$$f_\omega^*(p) \quad = \quad p F_\omega^*(p) + F_\omega(0),$$

whence, using equation (2.24)

$$(2.29) \qquad f_\omega^*(p) \quad = \quad \frac{f(\xi_1)}{\lambda_0}.$$

Let us suppose that $\mathcal{R}e(p) \geq \left\{\frac{\lambda_0+\mu}{1-\varepsilon}\right\}\varepsilon$. When $|z| = 1-\varepsilon$

$$|z| \quad > \quad \left|\frac{1}{\lambda_0+\mu+p}\{\mu + z f(z)\}\right|.$$

According to Theorem 2 of the appendix, ξ_1 is the unique root of the equation

$$z \quad = \quad \frac{1}{\lambda_0+\mu+p}\{\mu + z f(z)\}$$

within the region $|z| \leq 1-\varepsilon$.

The conditions of Lagrange's Theorem (see, for example, Whittaker and Watson [65,p133]) are satisfied, and we may expand any function of ξ_1 which is analytic within the region $|z| \leq 1-\varepsilon$ in a series. In particular

$$(2.30) \quad f(\xi_1) = \sum_{n=1}^{\infty} \left(\frac{1}{\lambda_0+\mu+p}\right)^n \frac{1}{n!} \frac{d^{n-1}}{dz^{n-1}} \left[\{\mu + z f(z)\}^n f'(z)\right]_{z=0}.$$

Since the series

$$e^{-(\lambda_0+\mu)t} \sum_{n=1}^{\infty} \frac{t^{n-1}}{\lambda_0 \, n!(n-1)!} \frac{d^{n-1}}{dz^{n-1}} \left[\{\mu + z f(z)\}^n f'(z)\right]_{z=0}$$

is uniformly convergent, we may take Laplace transforms term by term. The series we obtain in so doing is the right hand side of equation (2.30). Hence, using equation (2.29) we find

$$(2.31) \quad f_\omega(t) = e^{-(\lambda_0+\mu)t} \sum_{n=1}^{\infty} \frac{t^{n-1}}{\lambda_0 \, n!(n-1)!} \frac{d^{n-1}}{dz^{n-1}} \left[\{\mu + z f(z)\}^n f'(z)\right]_{z=0}.$$

Since $\frac{1}{n!}\{\mu + z f(z)\}^n$ is uniformly bounded for all z such that $|z| = 1-\varepsilon$ we may apply Cauchy's Integral Formula (see, for example, Copson [11,p69] and obtain

$$(2.32) \quad f_\omega(t) = e^{-(\lambda_0+\mu)t} \frac{1}{\lambda_0 t} \sum_{n=1}^{\infty} \frac{t^n}{n!} \frac{1}{2\pi i} \int_C \left[\frac{\mu}{z} + f(z)\right]^n f'(z)\, dz.$$

Again the uniform convergence of the right hand side of equation

(2.31) enables us to interchange the order of summation and integration and we find

$$(2.33)\quad f_w(t) = \frac{1}{2\pi i \lambda_0 t} e^{-(\lambda_0+\mu)t} \int_C \left[\exp\left\{\frac{\mu}{z}+f(z)\right\}t - 1\right] f'(z)\, dz.$$

Let us expand $\exp\left\{\frac{\mu}{z}+f(z)\right\}t$ as a Laurent series in z :

$$(2.34)\quad \exp\left\{\frac{\mu}{z}+f(z)\right\}t = \sum_{r=-\infty}^{\infty} z^r F_r(t).$$

Differentiating equation (2.34) with respect to z and comparing coefficients of z^{-1} we get

$$(2.35)\quad \sum_{r=1}^{\infty} r\lambda_r F_{-r}(t) = \mu F_1(t).$$

If we now evaluate the integral in the right hand side of equation (2.33), and use equations (2.34) and (2.35), we finally obtain

$$(2.36)\quad f_w(t) = \frac{\mu}{\lambda_0 t} e^{-(\lambda_0+\mu)t} F_1(t).$$

Applying exactly the same reasoning to the function

$$P_{i0}^{*}(p) = \frac{\xi_1^{i+1}}{\mu(1-\xi_1)},$$

we find that

$$(2.37)\quad P_{i0}(t) = \frac{1}{\mu t} e^{-(\lambda_0+\mu)t} \sum_{r=i+1}^{\infty} r F_{-r}(t).$$

Equations (2.36) and (2.37) are essentially the same as those which

Luchak [50] obtained for the probability density function of the busy period, and the probability that the system is empty at time t. Let us consider now three examples.

EXAMPLE 1: BATCHES OF CONSTANT SIZE

Let us suppose that $f(z) = \lambda z^m$. This corresponds to the case when customers arrive in batches of fixed size m. With this assumption equation (2.34) becomes

$$\exp\left(\lambda z^m + \frac{\mu}{z}\right)t = \sum_{r=-\infty}^{\infty} z^r F_r(t).$$

Now from the appendix, equation (a), we have

$$\exp\left(\lambda z^m + \frac{\mu}{z}\right)t = \sum_{r=-\infty}^{\infty} \left\{\left(\frac{\lambda}{\mu}\right)^{\frac{1}{m+1}} z\right\}^r I_{-r}^m(\alpha t), \text{ where } \alpha = 2(\lambda\mu^m)^{\frac{1}{m+1}}.$$

The functions $I_r^m(x)$ were first introduced by Luchak [49] and are defined in the appendix, equation (b). It follows that

$$F_r(t) = \left(\frac{\lambda}{\mu}\right)^{\frac{r}{m+1}} I_{-r}^m(\alpha t),$$

and so equations (2.36) and (2.37) become

$$(2.38) \qquad f_w(t) = \frac{\mu}{\lambda t}\left(\frac{\lambda}{\mu}\right)^{\frac{1}{m+1}} e^{-(\lambda+\mu)t} I_{-1}^m(\alpha t),$$

$$(2.39) \text{ and } \quad P_{i0}(t) = \frac{1}{\mu t} e^{-(\lambda+\mu)t} \sum_{r=i+1}^{\infty} r\left(\frac{\mu}{\lambda}\right)^{\frac{r}{m+1}} I_r^m(\alpha t).$$

Using equation (c) of the appendix, and rearranging, we get

(2.40) $$f_w(t) = \frac{m}{t}\left(\frac{\mu}{\lambda}\right)^{\frac{m}{m+1}} I_m^m(\alpha t)$$

(2.41) and $$P_{i0}(t) = e^{-(\lambda+\mu)t}\left\{\sum_{r=i}^{i+m}\left(\frac{\mu}{\lambda}\right)^{\frac{r}{m+1}} I_r^m(\alpha t) - \left(1-\frac{m\lambda}{\mu}\right)\sum_{r=i+m+1}^{\infty}\left(\frac{\mu}{\lambda}\right)^{\frac{r}{m+1}} I_r^m(\alpha t)\right\}.$$

Both of these results were obtained by Luchak [50].

We may find $P_{in}(t)$ for $n>0$ successively from the birth and death equations. For example, let us find $P_{i1}(t)$. From equation (2.4), with $\mu_s = \delta_{s1}\mu$ and $\lambda_0 = \lambda$, we have

$$P_{i1}(t) = \frac{1}{\mu}\frac{\partial}{\partial t}P_{i0}(t) + \frac{\lambda}{\mu}P_{i0}(t)$$

Putting in the expression for $P_{i0}(t)$ from equation (2.41) and using equation (e) of the appendix, we find

(2.42) $$P_{i1}(t) = e^{-(\lambda+\mu)t}\left\{\left(\frac{\mu}{\lambda}\right)^{\frac{i-1}{m+1}} I_{i-1}^m(\alpha t) + (1-m)\left(\frac{\mu}{\lambda}\right)^{\frac{i-1}{m+1}} I_{i+m}^m(\alpha t) + \sum_{r=i+1}^{\infty}\left(\frac{\mu}{\lambda}\right)^{\frac{r-1}{m+1}} I_{r+m}^m(\alpha t) + \frac{\lambda}{\mu}\left(1-\frac{m\lambda}{\mu}\right)\sum_{r=i+2m+1}^{\infty}\left(\frac{\mu}{\lambda}\right)^{\frac{r}{m+1}} I_r^m(\alpha t)\right\}.$$

When $m=1$, the batches are single customers, so we have the queue $M/M/1$. Since the function $I_r^m(x)$ reduces to the modified Bessel function $I_r(x)$ when $m=1$, (see appendix) equations (2.40) and (2.41) reduce to

(2.43) $$f_w(t) = \frac{1}{t}\sqrt{\frac{\mu}{\lambda}}\, e^{-(\lambda+\mu)t} I_1(2\sqrt{\lambda\mu}\,t),$$

(2.44) and $$P_{i0}(t) = e^{-(\lambda+\mu)t}\left\{\left(\frac{\mu}{\lambda}\right)^{\frac{i}{2}} I_i(2\sqrt{\lambda\mu}\,t) + \left(\frac{\mu}{\lambda}\right)^{\frac{i+1}{2}} I_{i+1}(2\sqrt{\lambda\mu}\,t) + \left(1-\frac{\lambda}{\mu}\right)\sum_{r=i+2}^{\infty}\left(\frac{\mu}{\lambda}\right)^{\frac{r}{2}} I_r(2\sqrt{\lambda\mu}\,t)\right\}.$$

It is not difficult to find a general formula for $P_{in}(t)$ in this case. Upon successive substitutions in the birth and death equations, starting with equation (2.44), we find

$$(2.45)\quad P_{in}(t) = e^{-(\lambda+\mu)t}\Big\{\left(\tfrac{\mu}{\lambda}\right)^{\frac{i-n}{2}} I_{i-n}(2\sqrt{\lambda\mu}\,t) + \left(\tfrac{\mu}{\lambda}\right)^{\frac{i-n+1}{2}} I_{i+n+1}(2\sqrt{\lambda\mu}\,t) + \left(1-\tfrac{\lambda}{\mu}\right)\left(\tfrac{\lambda}{\mu}\right)^{n} \sum_{r=i+n+2}^{\infty} \left(\tfrac{\mu}{\lambda}\right)^{\frac{r}{2}} I_r(2\sqrt{\lambda\mu}\,t)\Big\}.$$

This expression for $P_{in}(t)$ was first obtained by Ledermann and Reuter [48].

Suppose $\lambda \leq \mu$. Then the mean value of ω is given by

$$E(\omega) = \int_0^\infty t f_\omega(t)\,dt = \int_0^\infty e^{-(\lambda+\mu)t}\sqrt{\tfrac{\mu}{\lambda}}\, I_1(2\sqrt{\lambda\mu}\,t)\,dt.$$

Using Luke [57,p247 (7)] we find that

$$E(\omega) = \frac{1}{\mu-\lambda},$$

provided $\lambda < \mu$. If $\lambda = \mu$ the mean value is infinite.

By the previous remarks (see page 41), the above example corresponds to the queue $M/E_m/1$, where E_m denotes the distribution of a variable which is equal to the sum of m identical independent negative exponential variables. Such a variable is known as an Erlangian (or Gamma) variable, and has probability density function

$$f_\nu(x) = \frac{\mu}{\Gamma(m)}(\mu x)^{m-1} e^{-\mu x}.$$

Besides the example we have just considered Luchak worked out $P_{i0}(t)$

and $f_\omega(t)$ for one other example, namely the case when

$$f(z) = \lambda_1 z^{m_1} + \lambda_2 z^{m_2}.$$

Instead of considering this example, we will consider two further examples which have not been treated to date in the literature.

EXAMPLE 2: LOGARITHMIC DISTRIBUTION OF BATCH SIZE

Let us suppose that the batch size has probability mass function

$$f_r(n) = \frac{\lambda^n}{n}, \quad \text{where } 0 < \lambda < 1.$$

It follows that

$$f(z) = \log \frac{1}{1-\lambda z},$$

whence equation (2.33) becomes

$$(1-\lambda z)^{-t} e^{\frac{\mu}{z}t} = \sum_{r=-\infty}^{\infty} z^r F_r(t).$$

Comparing coefficients of z^r in the above, we have for $r \leq 0$

$$F_r(t) = \frac{(\mu t)^r}{r!} \, {}_1F_1(t; r+1; \lambda\mu t),$$

where ${}_1F_1$ is the confluent hypergeometric function, defined as

$$ {}_1F_1(t;r;x) = 1 + \frac{t}{r}x + \frac{t(t+1)}{2!\,r(r+1)}x^2 + \cdots\cdots\cdots, $$

(see, for example, Jeffries [30,p607]).

Thus equations (2.35) and (2.36) become

(2.46) $f_\omega(t) = \mu^2(\log\frac{1}{1-\lambda})^{-1}(1-\lambda)^{-t}e^{-\mu t}\,{}_1F_1(t;2;\lambda\mu t),$

(2.47) and $P_{i0}(t) = (1-\lambda)^t e^{-\mu t}\sum_{r=i}^{\infty}\frac{(\mu t)^r}{r!}\,{}_1F_1(t;r;\lambda\mu t).$

The condition that the busy period distribution be an "honest" distribution is, from equation (2.28), $(1-\lambda)^{-1} \leq \mu$.

EXAMPLE 3: GEOMETRIC DISTRIBUTION OF BATCH SIZE

As a final example we will consider the queue in which the size of the batch has probability distribution λ^r, where $0 < \lambda < 1$. In this case

$$ f(z) = \sum_{r=1}^{\infty}\lambda^r z^r = \frac{\lambda z}{1-\lambda z}, $$

and so the generating function for the $F_r(t)$ is $\exp\left(\frac{\lambda z}{1-\lambda z}+\frac{\mu}{z}\right)t$.
The denominator to the right hand side of equation (2.13) becomes

$$ \frac{-\lambda z}{1-\lambda z} + \left(\frac{\lambda}{1-\lambda}+\mu+p\right) - \frac{\mu}{z}, $$

whose zero inside the contour $C: |z| = 1-\varepsilon$ can be written down explicitly. Instead of finding $F_r(t)$ directly we will find $P_{in}(t)$

by picking out the coefficient of z^n in the right-hand side of equation (2.13) and inverting the resulting expression for $P_{in}^*(t)$.

Substituting $f(z) = \frac{\lambda z}{1-\lambda z}$ in equation (2.13) we get

$$(2.48) \quad \Pi^*(z, p) = \frac{(1-\lambda z)\{z^{i+1} - \mu(1-z) P_{i0}^*(p)\}}{-\lambda(1 - \frac{\lambda}{1-\lambda} + \mu + p) z^2 + (\frac{\lambda}{1-\lambda} + \mu + \lambda\mu + p) z - \mu}.$$

Let ξ_1 and ξ_2 be the zeros of the denominator of the right-hand side of equation (2.48). By Theorem 2 of the appendix there is one zero inside C. Let this zero be ξ_1. Putting equation (2.21) in equation (2.48) we get for $i = 0$

$$\Pi^*(z,p) = \frac{1-\lambda z}{\lambda \xi_2 (1 + \frac{\lambda}{1-\lambda} + \mu + p)(1 - \xi_1)} \left(1 - \frac{z}{\xi_2}\right)^{-1},$$

and comparing coefficients of z^n, we obtain

$$(2.49) \quad P_{0n}^*(p) = \frac{1 - \lambda \xi_2}{\lambda(1 + \frac{\lambda}{1-\lambda} + \mu + p)(1-\xi_1)\xi_2^{n+1}}, \text{ where } n \geq 1.$$

now

$$\xi_k = \left\{ \frac{\lambda}{1-\lambda} + \mu + \lambda\mu + p \pm \sqrt{\left(\frac{\lambda}{1-\lambda} + \mu + \lambda\mu + p\right)^2 - 4\lambda\mu\left(1 + \frac{\lambda}{1-\lambda} + \mu + p\right)} \right\} \Big/ 2\lambda\left(1 + \frac{\lambda}{1-\lambda} + \mu + p\right),$$

where $k=1$ corresponds to the negative square root. Substituting for ξ_1 and ξ_2 in equation (2.49) we have

$$(2.50) \quad P_{0n}^*(p) = 4\lambda^n \mu S^{-2}\left(1 + \frac{2}{S}\right)^{n-1}\left(1 - \frac{2\mu(1-\lambda)}{S}\right)^{-1}; \quad n \geq 1,$$

(2.51) where $S = \frac{\lambda}{1-\lambda} + \mu - \lambda\mu + p + \sqrt{\left(\frac{\lambda}{1-\lambda} + \mu - \lambda\mu + p\right)^2 - 4\lambda\mu}$.

We can expand the right hand side of equation (2.50) as a series in negative powers of S, since $\left|\frac{2\mu(1-\lambda)}{S}\right| < 1$ when $Re(p) > 0$, and so we have

$$(2.52) \quad P^*_{on}(p) = \lambda^n \left\{ \sum_{r=0}^{n-2} \sum_{k=0}^{r} \binom{n-1}{r-k} \mu^k (1-\lambda)^k S^{-(r+2)} + \left[1 + \frac{1}{\mu(1-\lambda)}\right]^{n-1} \sum_{r=n-1}^{\infty} \mu^r (1-\lambda)^r S^{-(r+2)} \right\}, \; n \geq 1.$$

Now S^{-r} is the Laplace transform of

$$\exp\left[-t\left(\frac{\lambda}{1-\lambda} + \mu - \lambda\mu + p\right]\right. \cdot (2\sqrt{\lambda\mu})^{-r} \frac{r}{t} I_r(2\sqrt{\lambda\mu}\, t)$$

(see, for example, Erdelyi [17] p240). Inverting each side of equation (2.51) term by term, we find

$$(2.53) \quad P_{on}(t) = \frac{\lambda}{\mu}^{n-1} \exp\left\{-t\left(\frac{\lambda}{1-\lambda} + \mu - \lambda\mu\right)\right\} \cdot \left[\sum_{r=0}^{n-2} \sum_{k=0}^{r} \binom{n-1}{r-k} \frac{\mu^k (1-\lambda)^k}{(\lambda\mu)^{r/2}} \frac{(r+2)}{t} I_{r+2}(2\sqrt{\lambda\mu}\, t) \right.$$
$$\left. + \left(1 + \frac{1}{\mu(1-\lambda)}\right)^{n-1} \sum_{r=n-1}^{\infty} \left(\frac{\mu(1-\lambda)}{\sqrt{\lambda\mu}}\right)^r \frac{(r+2)}{t} I_{r+2}(2\sqrt{\lambda\mu}\, t) \right], \; n \geq 1.$$

In a similar way we have by inverting equation (2.21)

$$(2.54) \quad P_{oo}(t) = \exp\left\{-t\left(\frac{\lambda}{1-\lambda} + \mu - \lambda\mu\right)\right\} \sum_{r=1}^{\infty} \frac{r}{\mu t} \left(\frac{\mu}{\lambda}\right)^{\frac{r}{2}} (1-\lambda)^{r-1} I_r(2\sqrt{\lambda\mu}\, t), \text{ when } i = 0.$$

Using the identity of the appendix (c), which reduces to

$$\frac{2r}{x} I_r(x) = I_{r-1}(x) - I_{r+1}(x)$$

when $n = 1$, and rearranging the terms in the right-hand side of equation (2,54), we obtain

$$(2.55)\quad P_{00}(t) = \exp\{-t(\tfrac{\lambda}{1-\lambda} + \mu - \lambda\mu)\}.\left[I_0(2\sqrt{\lambda\mu}t) + \rho^{-\frac{1}{2}} I_1(2\sqrt{\lambda\mu}t) + (1-\rho)\sum_{r=2}^{\infty} \rho^{-\frac{r}{2}} I_r(2\sqrt{\lambda\mu}t)\right],$$

where $\rho = \dfrac{\lambda}{\mu(1-\lambda)^2}$.

We will obtain the density function of the busy period for this example directly from equation (2.37). Putting $f(z) = \frac{\lambda z}{1-\lambda z}$ in equation (2.34) we get

$$\exp t\left\{\frac{\mu}{z} + \frac{\lambda z}{1-\lambda z}\right\} = \sum_{r=-\infty}^{\infty} z^r F_r(t),$$

and so in this case

$$(2.56)\qquad F_1(t) = e^{\lambda\mu t}\sqrt{\frac{\lambda}{\mu}}\, I_1(2\sqrt{\lambda\mu}t).$$

Substituting equation (2.56) into equation (2.36), we obtain

$$(2.57)\qquad f_w(t) = \frac{1}{\sqrt{\rho}\,t}\exp\{-\mu(1+\rho)t - \lambda\mu t\}\, I_1(2\sqrt{\lambda\mu}t),$$

where again we have put $\rho = \dfrac{\lambda}{(1-\lambda)^2}$.

It is interesting to note that equations (2.55) and (2.57), except for a factor $e^{-\lambda\mu t}$, take exactly the same form as the corresponding results for the queue with single arrivals and departures.

The condition that the busy period be an "honest" distribution is, from equation (2.28)

$$\frac{\lambda}{(1-\lambda)^2} \leq \mu.$$

Again using Luke [51, p247(7)] we find

$$E(\omega) = \frac{\rho}{\lambda(1-\rho)},$$

provided $\rho < 1$. If $\rho = 1$, the mean value of ω is infinite.

C. THE QUEUE WITH SINGLE ARRIVALS AND BATCH DEPARTURES.

If we impose the restriction $\lambda_r = \delta_{r,1}\lambda$, we have the queue with single arrivals and batch departures. Bailey [3] and Downton [15] have considered this problem for the more general queue with arbitrary service-time distribution, but their results are limited to the case $t \to \infty$. More recently Keilson [41] considered the behaviour of the queue for finite time, and obtained the same expression for the probability density function of ω as we shall obtain, in one particular case. These results were obtained independently of Keilson, who used quite different methods.

For the case of single departures we were able to express $P^*_{i0}(p)$ and $f^*_\omega(p)$ in terms of ξ_1, the only zero of the function

(2.58) $$-f(z) + (\lambda_0 + \mu_0 + p) - h(z)$$

within the contour $|z| = 1 - \varepsilon$, whenever $Re(p) \geq \delta$. For the case of batch departures the function given by equation (2.58) has m zeros in this region, and it would seem that Lagrange's expansion cannot be used. However when $f(z) = \lambda z$, we observe that the expression(2.58) has $m+1$ zeros in all, and so there is only one zero <u>outside</u> the contour $|z| = 1 - \varepsilon$, whenever $\mathcal{Re}(p) \geq \delta$. Consequently we will express $P_{i0}^*(p)$ and $f_\omega^*(p)$ in terms of this zero and apply Lagrange's Theorem as before.

<u>THEOREM 2.2</u>: There exist functions $g(z)$ and $G_n(z)$ such that

(2.59) $$f_\omega^*(p) = g\left(\frac{1}{\eta}\right),$$

(2.60) $$\text{and } P_{in}^*(p) = G_n\left(\frac{1}{\eta}\right), \quad n = 0, 1, \cdots, m-1,$$

where η is the unique zero of

$$-\lambda z + (\lambda + \mu_0 + p) - h(z)$$

provided $|z| \geq 1 - \varepsilon$ and $\mathcal{Re}(p) \geq \delta$.

<u>PROOF</u>: Since equations (2.15) are linear in $P_{in}^*(p)$, $0 \leq n < m$, and supposed independent, we may apply Cramer's rule. We have

(2.61) $$P_{in}^*(p) = \frac{|A_{n+1}|}{|A|}, \quad n = 0, 1, \cdots, m-1,$$

where $|A|$ is the determinant whose (i, j) element is

$$h(\xi_i) - \mu_0 \quad \text{if } j = 1,$$

$$\left\{(1-q)(h(\xi_i) - \mu_0) + q \sum_{s=j}^{m} \mu_s \left(\xi_i^{-s} - \xi_i^{1-j}\right)\right\} \xi_i^{j-1} \text{ otherwise},$$

and $|A_n|$ represents $|A|$ with element (r,n) replaced by ξ_r^i, $r = 1,2,\cdots,m$. Since interchange of ξ_i and ξ_j in each determinant corresponds to interchange of rows i and j, and thus reversal of sign, $|A|$ and $|A_n|$ are each alternating polynomials in the ξ_r, $r = 1,2,\cdots,m$. It follows from the properties of alternating polynomials (see, for example, Aitken [1,chVl]) that

$$|A_{n+1}| = \prod_{i<j}^{m} (\xi_j - \xi_i) s_{n+1},$$

$$\text{and } |A| = \prod_{i<j}^{m} (\xi_j - \xi_i) s,$$

where s and s_{n+1}, $n = 0,1,\cdots,m-1$ are symmetric polynomials in the ξ_r. Using equation (2.61) we have

$$P_{in}^*(p) = \frac{s_{n+1}}{s}.$$

Now using the properties of the zeros of polynomials (see, for example, Todhunter [63,p165]), any rational symmetric function of all but one of the zeros of a polynomial may be expressed as a function of the remaining zero only. Hence

$$P_{in}^{*}(p) = G_n\left(\frac{1}{\eta}\right),$$

where η is the zero of the polynomial

$$z^m\{-\lambda z + (\lambda + \mu_0 + p) - h(z)\}$$

which is outside the contour $|z| = 1 - \varepsilon$, provided $Re(p) \geq \delta$.

It follows by the same argument that

$$f_w^{*}(p) = g\left(\frac{1}{\eta}\right).$$

Let us suppose that it is possible to find a contour D , which lies wholly inside the contour

$$|z| = \frac{1}{1-\varepsilon}$$

and which encloses the origin, such that

(2.62) $\quad |z| > \left| \frac{1}{\lambda + \mu_0 + p}\{\lambda + z\, h(\tfrac{1}{z})\}\right|$ on D ;

(2.63) $\quad g(z)$ and $G_n(z)$ are analytic within D .

Let us also assume that

(2.64) $\quad g(0) = G_n(0) = 0, \text{ for } n = 0, 1, \cdots, m-1.$

We will justify these assumptions in the particular cases. By Lagrange's theorem, assumption (2.62) ensures that the equation

$$z = \frac{1}{\lambda+\mu_0+p}\left\{\lambda + z\,h\left(\frac{1}{z}\right)\right\}$$

has a unique root within D. This root is obviously $\frac{1}{\eta}$ since η is the only zero of the expression (2.58) in the region $|z| \geq 1-\varepsilon$, $Re(p) \geq \delta$.

Assumption (2.63) enables us to expand $g(\frac{1}{\eta})$ and $G_n(\frac{1}{\eta})$ in Lagrange series. Using equations (2.59) and (2.60) we find

$$f_w(p) = g(0) + \sum_{n=1}^{\infty}\left(\frac{1}{\lambda+\mu_0+p}\right)^n \frac{1}{n!}\,\frac{d^{n-1}}{dz^{n-1}}\left[g'(z)\left\{\lambda + z\,h\left(\tfrac{1}{z}\right)\right\}^n\right]_{z=0},$$

$$\text{and } P_{in}^*(p) = G_n(0) + \sum_{n=1}^{\infty}\left(\frac{1}{\lambda+\mu_0+p}\right)^n \frac{1}{n!}\,\frac{d^{n-1}}{dz^{n-1}}\left[G_n'(z)\left\{\lambda + z\,h\left(\tfrac{1}{z}\right)\right\}^n\right]_{z=0},$$

$$n = 0, 1, \dots, m-1.$$

Using the assumption (2.64), and proceeding as for equation (2.30) in section B, we obtain

$$\text{(2.65)} \quad f_w(t) = e^{-(\lambda+\mu_0)t}\,\frac{1}{2\pi i t}\int_D \left[\exp\left\{\frac{\lambda}{z} + h\left(\tfrac{1}{z}\right)\right\}t - 1\right] g'(z)\,dz,$$

$$\text{(2.66) and } P_{in}(t) = e^{-(\lambda+\mu_0)t}\,\frac{1}{2\pi i t}\int_D \left[\exp\left\{\frac{\lambda}{z} + h\left(\tfrac{1}{z}\right)\right\}t - 1\right] G_n'(z)\,dz, \quad 0 \leq n < m.$$

We will now proceed to find explicit expressions for $P_{in}(t)$ and $f_w(t)$ for some particular examples.

EXAMPLE 1. SERVER REMAINS IDLE UNTIL QUOTA IS REACHED $(q = 0)$

This example is the queue in which the server will serve m, and only m, customers simultaneously, and remains idle if the number of customers in the system is fewer than m. This corresponds to putting $q = 0$ in the birth and death equations. To begin with, let us work out the function $f_w(t)$.

Putting $q = 0$ in equation (2.16) we get

$$\hat{P}^*_{i0}(p) + \sum_{n=1}^{m-1} \xi_r^n P^*_{in}(p) = \frac{1}{p}\xi_r^i, \quad r = 1, 2, \cdots, m-1. \tag{2.67}$$

Now for single arrivals, the number in the system immediately after a busy period commences is m. Hence $i = m$, and from equation (1.37) we find

$$F_w(t) = \hat{P}_{m0}(t) + \sum_{n=1}^{m-1} \hat{P}_{mn}(t),$$

whence

$$f^*_w(p) = p \sum_{n=0}^{m-1} \hat{P}^*_{mn}(p), \tag{2.68}$$

after differentiating and taking Laplace transforms.
Now using Cramer's rule to solve equation (2.67) for $\hat{P}^*_{m0}(p)$, we have

$$\hat{P}^*_{mn}(p) = \frac{|A_{n+1}|}{|A|} \tag{2.69}$$

where $|A_{n+1}| = \begin{vmatrix} 1 & \xi_1 & \cdots & \xi_1^{n} & \frac{1}{p}\xi_1^{m} & \xi_1^{n+2} & \cdots & \xi_1^{m-1} \\ 1 & \xi_2 & \cdots & & & & & \\ \vdots & \vdots & & & \ddots & & & \\ 1 & \xi_m & & & & & & \xi_m^{m-1} \end{vmatrix}$ and $|A| = \begin{vmatrix} 1 & \xi_1 & \xi_1^2 & \cdots & \xi_1^{m-1} \\ 1 & \xi_2 & & & \\ \vdots & & & \ddots & \\ 1 & & & & \xi_m^{m-1} \end{vmatrix}$.

An appeal to the theory of alternants (see, for example, Aitken [1 ,chVl]) reveals that

$$|A_{n+1}| = \frac{1}{p}(-1)^{m-n+1} s_{m-n} |A|,$$

where s_{m-n} , is the (symmetric) sum of products of the ξ_r , taken $m-n$ at a time.

Using equations (2.68) and (2.69) we obtain

$$f_w^*(p) = \sum_{n=0}^{m-1} (-1)^{m-n+1} s_{m-n} = 1 - \prod_{i=1}^{m} (1-\xi_i).$$

Since $\xi_1, \xi_2, \cdots\cdots, \xi_m$ and η are the roots of the equation

$$-\lambda z^{m+1} + (\lambda + \mu_0 + p) z^m - z^m h(z) = 0,$$

we have

$$\prod_{i=1}^{m} (1-\xi_i) = \frac{p}{-\lambda(1-\eta)},$$

whence

(2.70) $$f_w^*(p) = 1 - \frac{p}{\lambda(\eta-1)} = \frac{\mu_0 - h(\eta)}{\lambda(\eta-1)}.$$

We have shown that in this case the function $g(z)$, introduced in equation (2.59), is $z(\mu_0 - k(\frac{1}{z}))/\lambda(1-z)$. This function is analytic within the region $|z| = 1$, and vanishes at $z = 0$. On the contour $|z| = 1 - \varepsilon$

$$\left| \lambda + z k(\tfrac{1}{z}) \right| \leq \lambda + \sum_{s=1}^{m} \mu_s z^{s+1} < \lambda + \mu_0 .$$

Thus assumption (2.62) is satisfied whenever $\mathcal{Re}(p) \geqslant 0$ and assumptions (2.63) and (2.64) are satisfied. Consequently equation (2.65) is valid, and in this case becomes

$$(2.71) \quad f_\omega(t) = \frac{1}{2\pi\lambda i t} e^{-(\lambda+\mu_0)t} \oint_D \left[\exp t\left\{ \frac{\lambda}{z} + k(\tfrac{1}{z}) \right\} - 1 \right] \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} j z^{j-1} dz.$$

Let us now write

$$(2.72) \quad \exp t\left\{ \frac{\lambda}{z} + k(\tfrac{1}{z}) \right\} = \sum_{n=-\infty}^{\infty} z^n H_n(t).$$

Then we may evaluate the integral in equation (2.71). We get

$$(2.73) \quad f_\omega(t) = \frac{1}{\lambda t} e^{-(\lambda+\mu_0)t} \sum_{s=1}^{m} s\left(\mu_0 - \sum_{j=1}^{s-1} \mu_j\right) H_{-s}(t).$$

We have not yet considered the condition for $F_\omega(t)$ to be a probability distribution function.

Using equations (2.26) and (2.68) we find

$$\lim_{t \to \infty} F_\omega(t) = \lim_{p \to 0+} f_\omega^*(p).$$

Hence from equation (2.70)

$$\lim_{t \to \infty} F_\omega(t) = \frac{1}{\lambda}\sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} \{\eta(0)\}^{-j}, \tag{2.74}$$

where $\eta(0)$ is the zero of

$$-\lambda z^{m+1} + (\lambda + \mu_0) z^m - z^m k(z)$$

which has greatest modulus.

<u>THEOREM 2.3</u>: Let $\eta(0)$ be the zero of $-\lambda z + (\lambda + \mu_0) - k(z)$ with the greatest modulus. $\eta(0)$ is real and if $\lambda \geq |k'(1)|$, $\eta(0) = 1$; otherwise $\eta(0) > 1$.

<u>PROOF</u>: Let us consider the function

$$\frac{-\lambda z^{m+1} + (\lambda + \mu_0) z^m - z^m k(z)}{1 - z} = \lambda z^m - \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} z^j .$$

If $\lambda \geq \sum_{s=1}^{m} s\mu_s$, the inequality

$$\left| \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} z^j \right| < |\lambda z^m|$$

is satisfied at all points on the circle $|z| = 1 - \varepsilon$, and so by Rouche's theorem the function $\lambda z^m - \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} z^j$ has m zeros within this region. Thus $\eta(0) = 1$. If $\lambda < \sum_{s=1}^{m} s\mu_s$, the curves $y = \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} x^{m-j}$ and $y = \lambda$ intersect for a unique x such that $0 \leq x < 1$.

THE UNIVERSITY OF TASMANIA

P.O. Box 252-C
HOBART
Telephone 2 7741

IN REPLY PLEASE QUOTE KRS/TC.

25th November 1964.

Professor P. A. Moran,
Institute of Advanced Studies,
Australian National University,
P.O. Box 4,
CANBERRA CITY, A.C.T.

Dear Professor Moran,

Mr. McNeil has discovered an error in his M.Sc. thesis and has asked that the proof of theorem 2.3 on page 65 be replaced by the enclosed statement. I should be grateful if you would incorporate this in the thesis already in your possession.

Yours faithfully,

(D. A. Kearney)
Registrar.

Encl.

Replace proof of theorem 2.3 (page 65) by:

"PROOF: Let us consider the function

$$\frac{-\lambda z^{m+1}+(\lambda+\mu_0)z^m - z^m \mathcal{L}(z)}{1-z} = \lambda z^m - \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} z^{m-j}.$$

If $\lambda \geq \sum_{s=1}^{m} s\mu_s$, the inequality

$$\left| \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} z^{m-j} \right| < \left| \lambda z^m \right|$$

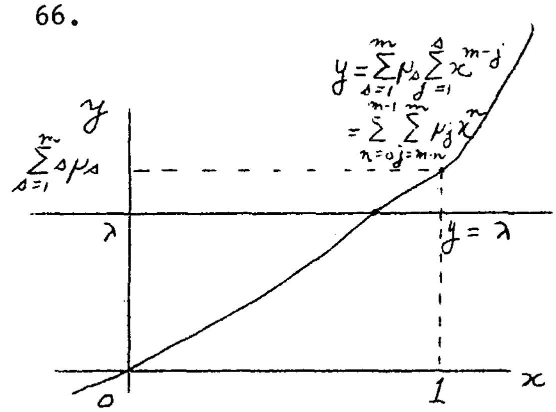
is satisfied at all points on the circle $|z| = 1-\varepsilon$, and so by Rouche's theorem the function $\lambda z^m - \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} z^{m-j}$ has m zeros within this region. Thus $\eta(0) = 1$.
If $\lambda < \sum_{s=1}^{m} s\mu_s$, then for sufficiently small $\varepsilon > 0$, which depends upon $\sum_{s=1}^{m} s\mu_s - \lambda$, it may be shown that

$$\lambda + \sum_{s=1}^{m} \mu_s (1-\varepsilon)^{s+1} < (\lambda+\mu_0)(1-\varepsilon).$$

Hence on $|z| = 1-\varepsilon$, $\left|\lambda + \sum_{s=1}^{m} \mu_s z^{s+1}\right| < |(\lambda+\mu_0)z|$, and so by Rouche's theorem the function $\lambda - (\lambda+\mu_0)z + \sum_{s=1}^{m} \mu_s z^{s+1}$ has exactly one zero in $|z| < 1-\varepsilon$. It follows that the function $\lambda z^{m+1} - (\lambda+\mu_0)z^m + \sum_{s=1}^{m} \mu_s z^{m-s}$ has exactly one zero in $|z| > \frac{1}{1-\varepsilon}$. Furthermore, this zero is real since the curves $y = \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} x^j$ and $y = \lambda$ intersect for $0 < x < 1$, which shows that

$$\lambda z^m - \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} z^{m-j}$$

has a real zero, $\eta(0)$, which is greater than unity. This completes the proof."



Consequently the equation

$$\lambda z^{-m} = \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} z^{-j}$$

has a unique root in $[0, 1)$, and so the function $\lambda z^m - \sum_{s=1}^{m} \mu_s \sum_{j=1}^{s} z^j$ has a unique real zero greater than unity. Hence $\eta(0) > 1$.

Now if $\eta(0) > 1$ it follows from equation (2.74) that

$$\lim_{t \to \infty} F_\omega(t) = \frac{1}{\lambda} \,|k'(1)| \,.$$

If, on the other hand, $\eta(0) = 1$, from equations (2.70) and (2.26) we have $\lim_{t \to \infty} F_\omega(t) = 1$.

We have shown that $F_\omega(t)$ is a probability distribution function if and only if $\lambda \leq |k'(1)|$. If $\lambda > |k'(1)|$ there is a probability $1 - |k'(1)|/\lambda$ that a busy period never ends. It may be shown, as in section B (see p45) that λ is the mean arrival rate and $|k'(1)|$ is the mean rate of departures from service. Thus the condition for the busy periods to be of finite length is that the mean arrival rate does not exceed the mean rate of departures from service. This is what we would expect.

In the particular case $\mu_s = \delta_{sm}\mu$, that is when departures occur in batches of fixed size, equation (2.73) reduces to

$$(2.75) \quad f_w(t) = \frac{\mu}{\lambda t} e^{-(\lambda+\mu)t} \sum_{s=1}^{m} s\left(\frac{\lambda}{\mu}\right)^{\frac{s}{m+1}} I_s^m\left(2(\lambda^m\mu)^{\frac{1}{m+1}} t\right),$$

since, in this case, we obtain from equation (2.72)

$$H_n(t) = \left(\frac{\lambda}{\mu}\right)^{-\frac{n}{m+1}} I_{-n}^m\left(2(\lambda^m\mu)^{\frac{1}{m+1}} t\right)$$

(see equation (a) of the appendix).

As is expected, when $m = 1$, equation (2.75) reduces to the well known formula for the queue $M/M/1$:

$$f_w(t) = e^{-(\lambda+\mu)t} \sqrt{\frac{\mu}{\lambda}} \frac{1}{t} I_1(2\sqrt{\lambda\mu}\, t).$$

Expressions for $P_{in}(t)$ may be found in a similar manner. However they are complicated, even in the very simple case $m = 2$, $i = 0$, $\mu_s = \delta_{s2}\mu$. For this particular example we have, from equation (2.61)

$$P_{00}^*(p) = \frac{-\xi_1\xi_2(1-\xi_1\xi_2)}{\mu(1-\xi_1^2)(1-\xi_2^2)} \quad \text{and} \quad P_{01}^*(p) = \frac{\xi_1 + \xi_2}{\mu(1-\xi_1^2)(1-\xi_2^2)} .$$

Substituting the relations $\xi_1 \xi_2 = -\frac{\mu}{\lambda\eta}$ and $\xi_1 + \xi_2 = \frac{\mu}{\lambda\eta^2}$, and using equation (2.60), we get

$$(2.76)\quad G_0(z) = \frac{z(\lambda + \mu z)}{(\lambda - \mu z - \mu z^2)(\lambda - \mu z + \mu z^2)}, \text{ and } G_1(z) = \frac{\lambda z^2}{(\lambda - \mu z - \mu z^2)(\lambda - \mu z + \mu z^2)}.$$

$G_0(z)$ and $G_1(z)$ have singularities at $\frac{1}{2}(-1 \pm \sqrt{1 + 4\frac{\lambda}{\mu}})$ and $\frac{1}{2}(1 \pm \sqrt{1 - 4\frac{\lambda}{\mu}})$. Of these $\alpha = \frac{1}{2}(-1 + \sqrt{1 + 4\frac{\lambda}{\mu}})$ has smallest modulus. Let us choose D to be the contour $|z| = \alpha - \varepsilon$, where $\varepsilon > 0$ is small. Then on D

$$(2.77)\quad |(\lambda + \mu z^3)(\lambda + \mu_0 + p)^{-1}| \leq \{\lambda + \mu(\alpha - \varepsilon)^3\}\{\lambda + \mu_0 + Re(p)\}^{-1}.$$

Now since α is the greater zero of $\lambda - \mu z - \mu z^2$, we have

$$\lambda - \mu(\alpha - \varepsilon) - \mu(\alpha - \varepsilon)^2 < \delta$$

where $\delta > 0$ is arbitrarily **small**, provided ε is sufficiently small. Multiplying through by $1 - (\alpha - \varepsilon)$, which is positive, we get

$$\lambda - (\lambda + \mu)(\alpha - \varepsilon) + \mu(\alpha - \varepsilon)^3 < \delta(1 - \alpha + \varepsilon),$$

whence $\quad \lambda + \mu(\alpha - \varepsilon)^3 < \{\lambda + \mu + \delta(\frac{1 - \alpha + \varepsilon}{\alpha - \varepsilon})\}(\alpha - \varepsilon).$

Using the inequality (2.77) we find that

$$|(\lambda + \mu z^3)(\lambda + \mu_0 + p)^{-1}| < |z|$$

on $\mathcal{D}$, provided $Re(p) > \delta\left(\frac{1-\alpha+\varepsilon}{\alpha-\varepsilon}\right)$.

We have thus shown that the assumptions (2.62), (2.63) and (2.64) are satisfied in this case, and so equation (2.66) is valid. We obtain, using equations (2.72) and (2,76)

$$P_{00}(t) = \frac{1}{t} e^{-(\lambda+\mu)t} \sum_{r=1}^{\infty} a_r \left(\frac{\lambda}{\mu}\right)^{\frac{r}{3}} I_r^2 (2\lambda^{\frac{2}{3}}\mu^{\frac{1}{3}}t),$$

$$\text{and} \quad P_{01}(t) = \frac{1}{t} e^{-(\lambda+\mu)t} \sum_{r=1}^{\infty} b_r \left(\frac{\lambda}{\mu}\right)^{\frac{r}{3}} I_r^2 (2\lambda^{\frac{2}{3}}\mu^{\frac{1}{3}}t),$$

where a_r and b_r are the coefficients of z^r in $G_0'(z)$ and $G_1'(z)$ respectively.

EXAMPLE 2: THE CASE WHEN THE SERVER IS BUSY UNLESS THE SYSTEM IS EMPTY

The queue we are about to consider is the case in which the customers are served even when the quota is not reached. This corresponds to putting $q = 1$ in the birth and death equations. We will only consider the case $\mu_s = \delta_{sm}\mu$, which corresponds to the queue in which all the customers in service depart at the end of a service period. Instead of proceeding as we did for the previous example, however, we will find $P_{in}(t)$ and $f_\omega(t)$ by a quite different method. This method was first applied to the queue $M/M/1$ by Cox and Smith [12,p61], and later to queues with time-dependent parameters by Keilson [38]. We proceed as follows:

Putting $q = 1, \lambda_r = \delta_{r1}\lambda$ and $\mu_s = \delta_{sm}\mu$ in equations (2.2), (2.3) and (2.4) we obtain

$$\text{(2.78)} \qquad \frac{\partial}{\partial t} P_{1n}(t) = \lambda P_{1,n-1}(t) - (\lambda+\mu) P_{1n}(t) + \mu P_{1,n+m}(t), \; n \geq 1;$$

$$\text{(2.79) and} \qquad \frac{\partial}{\partial t} P_{10}(t) = -\lambda P_{10}(t) + \mu \sum_{s=1}^{m} P_{1s}(t).$$

Because of the impossibility of there being a negative number of customers in the system, equation (2.79) is different from the general form of equation (2.78). But let us suppose, for the moment, that equation (2.78) holds for all integral values of n. Then if we can find a solution to this extended system which satisfies the additional requirement that

$$\text{(2.80)} \qquad \frac{\lambda}{\mu} P_{1,-1}(t) = \sum_{s=0}^{m-1} P_{1s}(t),$$

then equations (2.78) and (2.79) will be satisfied.

Let us multiply equation (2.78) by z^n and sum over all integral values of n. Then the generating function

$$\Pi(z,t) = \sum_{n=0}^{\infty} z^n P_{1n}(t)$$

satisfies the partial differential equation

$$\frac{\partial}{\partial t} \Pi(z,t) = \{\lambda z - (\lambda+\mu) + \mu z^{-m}\} \Pi(z,t),$$

the most general solution to which is

$$\text{(2.81)} \qquad \Pi(z,t) = \phi(z) \exp\left[-(\lambda+\mu)t + (\lambda z + \mu z^{-m})t\right],$$

where ϕ is an arbitrary function of z .

Now using equation (a) of the appendix, equation (2.81) may be written

$$(2.82) \qquad \Pi(z,t) = \phi(z)\, e^{-(\lambda+\mu)t} \sum_{r=-\infty}^{\infty} z^r \left(\frac{\lambda}{\mu}\right)^{\frac{r}{m+1}} I_r^m(\beta t),$$

where $\beta = 2(\lambda^m \mu)^{\frac{1}{m+1}}$. If we now replace $\phi(z)$ by an expansion $\sum_{r=-\infty}^{\infty} a_r z^r$, and pick out coefficients of z^n in each side of equation (2.82), we find

$$(2.83) \qquad P_{in}(t) = e^{-(\lambda+\mu)t} \sum_{r=-\infty}^{\infty} a_r \left(\frac{\lambda}{\mu}\right)^{\frac{n+r}{m+1}} I_{n+r}^m(\beta t).$$

Since $P_{in}(0) = \delta_{in}$, from equation (1.34) and $I_n^m(0) = \delta_{n0}$ (see equation (f), appendix) we have, putting $t = 0$ in equation (2.83)

$$a_{-n} = \delta_{in} \qquad (n = 0, 1, 2, \ldots\ldots).$$

Hence

$$(2.84) \qquad P_{in}(t) = e^{-(\lambda+\mu)t} \left\{ \left(\frac{\lambda}{\mu}\right)^{\frac{n-i}{m+1}} I_{n-i}^m(\beta t) + \sum_{r=1}^{\infty} a_r \left(\frac{\lambda}{\mu}\right)^{\frac{n+r}{m+1}} I_{n+r}^m(\beta t) \right\}.$$

Now we will use equation (2.80) to determine the remaining a_r .

Putting equation (2.84) in equation (2.80) and considering the case $i = 0$ only, we find

$$(2.85) \qquad \left(\frac{\lambda}{\mu}\right)^{\frac{m}{m+1}} I_{-1}^m(\beta t) + \sum_{r=1}^{\infty} a_r \left(\frac{\lambda}{\mu}\right)^{\frac{m+r}{m+1}} I_{r-1}^m(\beta t) \equiv I_0^m(\beta t) + \sum_{r=1}^{\infty} a_r \left(\frac{\lambda}{\mu}\right)^{\frac{r}{m+1}} I_r^m(\beta t)$$
$$+ \cdots\cdots + \left(\frac{\lambda}{\mu}\right)^{\frac{m-1}{m+1}} I_{m-1}^m(\beta t) + \sum_{r=1}^{\infty} a_r \left(\frac{\lambda}{\mu}\right)^{\frac{r+m-1}{m+1}} I_{r+m-1}^m(\beta t).$$

Using the relation $I_{-1}^m(x) = m I_m^m(x)$, (see the appendix, equation (d)), and comparing the coefficients of $I_n^m(\beta t)$, (which form a linearly independent set) in equation (2.85), we get

$$\begin{aligned}
\frac{\lambda}{\mu} a_1 &= 1, \\
\frac{\lambda}{\mu} a_{n+1} &= a_n + a_{n-1} + \cdots + a_1 + 1, \quad 0 < n < m, \\
m + \frac{\lambda}{\mu} a_{m+1} &= a_m + a_{m-1} + \cdots + a_1 + 1, \\
\frac{\lambda}{\mu} a_{n+1} &= a_n + a_{n-1} + \cdots + a_{n-m+1}, \quad n < m.
\end{aligned}$$

Solving these equations recursively for the a_n we obtain

$$\begin{aligned}
a_n &= \frac{\mu}{\lambda}\left(\frac{\mu}{\lambda}+1\right)^{n-1}, \quad n \leq m, \\
a_{m+1} &= \frac{\mu}{\lambda}\left\{\left(\frac{\mu}{\lambda}+1\right)^m - m\right\}, \\
a_{m+j} &= \frac{\mu}{\lambda}\left(a_{m+j-1} + a_{m+j-2} + \cdots + a_j\right), \quad j > 1.
\end{aligned}$$

This determines the solution for $P_{0n}(t)$, namely

$$\text{(2.86)} \quad P_{0n}(t) = e^{-(\lambda+\mu)t}\left\{\left(\frac{\lambda}{\mu}\right)^{\frac{n}{m+1}} I_n^m(\beta t) + \frac{\mu}{\lambda}\sum_{r=1}^{m}\left(1+\frac{\mu}{\lambda}\right)^{r-1}\left(\frac{\lambda}{\mu}\right)^{\frac{n+r}{m+1}} I_{n+r}^m(\beta t) + \right.$$

$$\left. + \left[\left(1+\frac{\mu}{\lambda}\right)^m - m\right]\left(\frac{\lambda}{\mu}\right)^{\frac{n}{m+1}} I_{n+m+1}^m(\beta t) + \sum_{r=m+2}^{\infty} a_r \left(\frac{\lambda}{\mu}\right)^{\frac{n+r}{m+1}} I_{n+r}^m(\beta t)\right\}.$$

where $\beta = 2(\lambda^m\mu)^{\frac{1}{m+1}}$ and the a_r are given by the above recurrence relations.

To find $f_\omega(t)$, we solve the birth and death equations for the modified process, which are, from equations (2.5), (2.6) and (2.7)

$$(2.87) \quad \frac{\partial}{\partial t}\hat{P}_{in}(t) = \lambda \hat{P}_{i,n-1}(t) - (\lambda+\mu)\hat{P}_{in}(t) + \mu \hat{P}_{i,n+m}(t), \ n>1;$$

$$(2.88) \quad \frac{\partial}{\partial t}\hat{P}_{i1}(t) = -(\lambda+\mu)\hat{P}_{i1}(t) + \mu \hat{P}_{i,m+1}(t),$$

$$(2.89) \text{ and } \frac{\partial}{\partial t}\hat{P}_{i0}(t) = \mu \sum_{n=1}^{m} \hat{P}_{in}(t).$$

Since the server is busy unless the queue is empty, we have $i = 1$. Solving equations (2.87) and (2,88) in exactly the same way as for equations (2.78) and (2.79) we get

$$\hat{P}_{1n}(t) = e^{-(\lambda+\mu)t} \left(\frac{\lambda}{\mu}\right)^{\frac{n-1}{m+1}} \left\{ I_{n-1}^{m}(\beta t) - m I_{n+m}^{m}(\beta t) \right\},$$

whence, from equation (2.89), we have

$$(2.90) \quad \frac{\partial}{\partial t}\hat{P}_{10}(t) = \mu e^{-(\lambda+\mu)t} \sum_{n=1}^{m} \left(\frac{\lambda}{\mu}\right)^{\frac{n-1}{m+1}} \left\{ I_{n-1}^{m}(\beta t) - m I_{m+n}^{m}(\beta t) \right\}.$$

Now from equation (1.37), in this case $f_\omega(t) = \frac{\partial}{\partial t}\hat{P}_{10}(t)$. Thus, using the equation (c) of the appendix to simplify equation (2.90) we obtain

$$(2.91) \quad f_\omega(t) = \frac{\mu}{\lambda t} e^{-(\lambda+\mu)t} \sum_{n=1}^{m} n \left(\frac{\lambda}{\mu}\right)^{\frac{n}{m+1}} I_n^{m}(\beta t).$$

It is interesting to note that equation (2.91) takes exactly the same form as equation (2.75), for the density function of the busy period in the corresponding case for $q = 0$. Thus the condition for $F_\omega(t)$ to be a probability distribution function will be the same, namely $\lambda \leq |k'(1)|$. Equation (2.91) was obtained by Keilson [41], using a different method.

CHAPTER 3.

QUEUES WITH A STATE-DEPENDENT ARRIVAL PROCESS

If we confine our queue to that in which the arrival and departure processes are simple (see chapter 1,p30) we have what is known as the queue with state-dependent parameters. These queues arise when the probability of an arrival or departure in an interval $(t, t+\delta t]$ depends upon the number of customers in the system. To quote a realistic example, the sight of a large queue may discourage further customers from joining it.

Putting $\lambda_{sn} = \delta_{s1}\lambda_n$, $\mu_{sn} = \delta_{s1}\mu_n$ in equations (1.31), (1.32) and (1.33) we find

(3.1) $$\frac{\partial}{\partial t} P_{in}(t) = \lambda_{n-1} P_{i,n-1}(t) - (\lambda_n + \mu_n) P_{in}(t) + \mu_{n+1} P_{i,n+1}(t), \; n \geq 1,$$

(3.2) and $$\frac{\partial}{\partial t} P_{i0}(t) = -\lambda_0 P_{i0}(t) + \mu_1 P_{i1}(t).$$

These equations have generally been called "the birth and death equations" and have been studied extensively in the literature. They first arose in the study of population growth, where the λ_n and μ_n are linear functions of n. This particular problem was first considered, for the case $\lambda_n = \lambda n$, $\mu_n = \mu n$ only, by Kendall [42], who found the $P_{in}(t)$ by generating function techniques. The first investigation into the problem of solving equations (3.1) and (3.2) for general λ_n and μ_n, was conducted by Ledermann and Reuter [48] who used spectral theory. They obtained an integral representation for the $P_{in}(t)$ by solving first the associated problem obtained by limiting the number of customers in the system to N . By allowing N to tend to infinity, conditions were obtained for the solution to converge. While various special cases of the above system of equations have been investigated by Bather [4], Foster [23], Keilson [38], Sack [60], Heathcote and Moyal [27], Moran [54] and others, the most important contribution by far has been made by Karlin and McGregor [31], [32], [33], [34], and [35]. Their first paper [31] deals with the existence, uniqueness and properties of the functions $P_{in}(t)$ which satisfy equations (3.1), (3.2) and the initial condition

$$(3.3) \qquad P_{in}(0) = \delta_{in}$$

The second paper of Karlin and McGregor [32] involves a classification of all "birth and death processes" defined by equations (3.1) and (3.2) in terms of certain ergodic properties of the process, such as whether or not the number in the system is almost surely unbounded as $t \to \infty$. The remaining three papers [33], [34] and [35] are concerned with applications to particular functions λ_n and μ_n . One paper [33] treats the case for a queue with m identical servers, while in another [34] the λ_n and μ_n are each linear functions of n. In each case the $P_{in}(t)$ are expressed in terms of an integral, and the general method is quite different to the approach for batch queues, in that it involves matrices rather than generating functions.

The equations (1.40), (1.41) and (1.42) for the modified process, from which we obtain the probability density function of the busy periods, become in this case

$$(3.4) \qquad \frac{\partial}{\partial t}\hat{P}_{in}(t) = \lambda_{n-1}\hat{P}_{i,n-1}(t) - (\lambda_n+\mu_n)\hat{P}_{in}(t) + \mu_{n+1}\hat{P}_{i,n+1}(t),\ n>1,$$

$$(3.5) \qquad \frac{\partial}{\partial t}\hat{P}_{i1}(t) = -(\lambda_1+\mu_1)\hat{P}_{i1}(t) + \mu_2\hat{P}_{i2}(t),$$

$$(3.6) \text{ and } \frac{\partial}{\partial t}\hat{P}_{i0}(t) = \mu_1\hat{P}_{i1}(t).$$

If we ignore the state $N(t)=0$, equations (3.4) and (3.5) take almost the same form as equations (3.1) and (3.2). If equation (3.2) was replaced by

$$(3.7) \qquad \frac{\partial}{\partial t}P_{i0}(t) = -(\lambda_0+\mu_0)P_{i0}(t) + \mu_1 P_{i1}(t),$$

and "state n" was relabled "state $n-1$" in equations (3.4) and (3.5), the two systems would be identical. Karlin and McGregor actually investigated the more general birth and death equations (3.1) and (3.7). The case $\mu_0=0$ gives the original problem.

This chapter will largely be an exposition of the results and methods of Karlin and McGregor. It is not possible, of course, to give a complete exposition. Consequently we shall work through the case $\mu_0=0$ only. We shall proceed as follows:

In section A, we shall construct an integral representation for the $P_{in}(t)$, in a purely formal way. In section B, we shall prove that such a representation exists. In section C, we shall find a necessary and sufficient condition for the uniqueness of the integral representation. The remaining section will comprise a consideration of particular examples, including a new one.

A. FORMAL CONSTRUCTION OF INTEGRAL REPRESENTATION

Equations (3.1), (3.2) and (3.3) can be written in the matrix form

(3.8) $$P'(t) = P(t)\,A,$$

(3.9) $$\text{and } P(0) = I,$$

where $P(t)$ is the infinite matrix $\{P_{in}(t)\}$, $i, n = 0, 1, 2, \cdots$, and A is the infinite matrix

$$A = \begin{vmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \cdots \\ 0 & 0 & \mu_2 & -(\lambda_3+\mu_3) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$

Let us define a sequence of polynomials $\{Q_n(x)\}$ by the recurrence relation

(3.10) $$-x\,Q_n(x) = \mu_n Q_{n-1}(x) - (\lambda_n+\mu_n) Q_n(x) + \lambda_n Q_{n+1}(x),$$

where $Q_{-1}(x) \equiv 0$ and $Q_0(x) \equiv 1$. This may be written more compactly in the matrix form

(3.11) $$-x\,Q(x) = A\,Q(x),$$

where $Q(x)$ is the infinite column vector $\{Q_n(x)\}$. Let us write

$$(3.12) \qquad f_i(x,t) = \sum_{n=0}^{\infty} P_{in}(t)\, Q_n(x).$$

If $f(x,t)$ denotes the infinite column vector $\{f_i(x,t)\}$, we have

$$f(x,t) = P(t)\, Q(x),$$

whence upon using equation (3.8)

$$\frac{\partial}{\partial t} f(x,t) = P'(t)\, Q(x) = P(t)\, A\, Q(x).$$

Now using equation (3.11), we find

$$\frac{\partial}{\partial t} f(x,t) = -x\, P(t)\, Q(x) = -x f(x,t).$$

The solution to this partial differential equation is

$$f(x,t) = e^{-xt} Q(x),$$

using the initial condition $f(x,0) = Q(x)$, which follows from equations (3.9) and (3.12). Hence the functions $f_i(x,t)$ are given by

$$(3.13) \qquad f_i(x,t) = e^{-xt} Q_i(x).$$

Now let us suppose that the system $\{Q_n(x)\}$ is orthogonal with respect to a probability distribution $\psi(x)$ on the interval $[0,\infty)$. Then according to equation (3.12), $P_{in}(t)$ is the n^{th} Fourier coefficient in the expansion of $f_i(x,t)$. It follows that

$$P_{in}(t) = \pi_n \int_0^{\infty} f_i(x,t)\, Q_n(x)\, d\psi(x),$$

where $\frac{1}{\pi_n} = \int_0^\infty Q_n^2(x)\, d\psi(x).$

Using equation (3.13) this becomes

(3.14) $$P_{in}(t) = \pi_n \int_0^\infty e^{-xt} Q_i(x)\, Q_n(x)\, d\psi(x).$$

B. PROOF OF THE EXISTENCE OF THE INTEGRAL REPRESENTATION

In order to justify the above formal construction we shall show that (1) there is a non-decreasing function $\psi(x)$ with respect to which the system $\{Q_n(x)\}$ is orthogonal on the interval $[0, \infty)$, and

(2) the integral representation (3.14) satisfies equations (3.1), (3.2) and (3.3).

(1) EXISTENCE OF $\psi(x)$:

If $\psi(x)$ is to have total mass unity, we must have $\pi_0 = 1$. It follows from the recurrence relations (3.10) that $Q_n(x)$ is a polynomial of exact degree n, and that the coefficient of x^n is $(-1)^n/\lambda_0\lambda_1\cdots\cdots\lambda_{n-1}$.

We can find the moments of $\psi(x)$,

$$c_n = \int_0^\infty x^n\, d\psi(x), \quad n = 0, 1, 2, \cdots\cdots,$$

recursively by solving the equations

$$\int_0^\infty Q_0(x)\, d\psi(x) = 1, \quad \int_0^\infty Q_n(x)\, d\psi(x) = 0, \quad n = 1, 2, \cdots .$$

The problem of finding a distribution function when its moments are given is known as "the problem of moments". If the interval on which the measure is defined is the non-negative real axis, then the problem is called "the Stieltjes moment problem". Shohat and Tamarkin [61,p6] proved that the Stieltjes moment problem has at least one solution, which is not supported by a finite set of points, if and only if all the determinants

$$\Delta_n = \begin{vmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n} \end{vmatrix} \quad \text{and} \quad \Delta_n^{(1)} = \begin{vmatrix} c_1 & c_2 & \cdots & c_{n+1} \\ c_2 & c_3 & \cdots & c_{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n+1} & c_{n+2} & \cdots & c_{2n+1} \end{vmatrix}, \quad n = 0, 1, \cdots,$$

are strictly positive. Let us now show that this condition is satisfied for the system $\{Q_n(x)\}$.

From the recurrence relations (3.10) it follows that $x^k Q_n(x)$ may be expressed in the form

$$x^k Q_n(x) = \sum_{r=0}^{2k} \alpha_r Q_{n-k+r}(x), \quad \text{provided } 0 \le k \le n,$$

where α_r, $r = 0, 1, \ldots, 2k$ are constants.

Since $\int_0^\infty Q_m(x)\, d\psi(x) = 0$ for all $m > 0$, it follows that

(3.15) $$\int_0^\infty x^k Q_n(x)\, d\psi(x) = 0, \quad \text{provided } 0 \le k < n.$$

Also from equation (3.10) we have

$$(-x)^n Q_n(x) = (-x)^{n-1}\{\mu_n Q_{n-1}(x) - (\lambda_n + \mu_n) Q_n(x) + \lambda_n Q_{n+1}(x)\}.$$

Using equation (3.15) we obtain

$$\int_0^\infty (-x)^n Q_n(x)\,d\psi(x) = \mu_n \int_0^\infty (-x)^{n-1} Q_{n-1}(x)\,d\psi(x),$$

(3.16) whence $\int_0^\infty (-x)^n Q_n(x)\,d\psi(x) = \mu_1\mu_2\cdots\mu_n$.

Let us now consider the n^{th} degree polynomial defined by the determinant

$$D_n(x) = \begin{vmatrix} c_0 & c_1 & \cdots & c_{n-1} & 1 \\ c_1 & c_2 & \cdots & c_n & x \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n-1} & x^n \end{vmatrix}, \quad n \geq 1.$$

Using the definition of the moments c_n it is clear that

$$\int_0^\infty x^k D_n(x)\,d\psi(x) = \begin{vmatrix} c_0 & c_1 & \cdots & c_{n-1} & c_k \\ c_1 & c_2 & \cdots & c_n & c_{k+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n-1} & c_{2n} \end{vmatrix}$$

Consequently

(3.17) $$\int_0^\infty x^k D_n(x)\,d\psi(x) = \begin{cases} 0 & \text{if } 0 \leq k < n \\ \Delta_n & \text{if } k = n. \end{cases}$$

Hence, using equation (3.15), $D_n(x)$ is a constant multiple of $Q_n(x)$. (since putting $D_n(x) = \sum_{i=0}^{n} \alpha_i Q_i(x)$, and multiplying each side by x^k $(0 \leq k \leq n)$ and integrating from 0 to ∞ with respect to $\psi(x)$ all the α_i vanish except α_n.)

The coefficient of x^n in $D_n(x)$ is Δ_{n-1}, and since the coefficient of x^n in $Q_n(x)$ is $(-1)^n/\lambda_0\lambda_1\cdots\lambda_{n-1}$, we may write

$$D_n(x) = (-1)^n(\lambda_0\lambda_1\cdots\lambda_{n-1})\,\Delta_{n-1}\,Q_n(x). \tag{3.18}$$

If we now multiply equation (3.18) by x^n and integrate with respect to $\psi(x)$ we find, using equations (3.16) and (3.17),

$$\Delta_n = (\lambda_0\lambda_1\cdots\lambda_{n-1})(\mu_1\mu_2\cdots\mu_n)\,\Delta_{n-1}$$

Since $\Delta_0 = c_0 = 1$, it follows that $\Delta_n > 0$ for all n.
If we put $x = 0$ in equation (3.10) we observe that

$$Q_n(0) = 1, \text{ for all } n. \tag{3.19}$$

Now upon using equation (3.18) $\Delta_n^{(1)} = (-1)^{n+1}D_{n+1}(0) = (\lambda_0\lambda_1\cdots\lambda_n)\Delta_n Q_{n+1}(0)$.
Thus $\Delta_n^{(1)} > 0$ for all n.

We have shown that there exists at least one probability distribution function $\psi(x)$, defined on the interval $[0,\infty)$ and not supported by a finite set of points, whose moments are the c_n. Since the leading coefficient in $Q_n(x)$ is $(-1)^n/\lambda_0\lambda_1\cdots\lambda_{n-1}$, it follows from equations (3.15) and (3.16) that

$$\int_0^\infty Q_n(x)\,Q_m(x)\,d\psi(x) = \delta_{mn}\,\frac{\mu_1\mu_2\cdots\mu_n}{\lambda_0\lambda_1\cdots\lambda_{n-1}}. \tag{3.20}$$

Hence the system $\{Q_n(x)\}$ is orthogonal with respect to $\psi(x)$ and

$$\pi_n = \frac{\lambda_0\lambda_1\cdots\lambda_{n-1}}{\mu_1\mu_2\cdots\mu_n} \text{ for } n \geq 1, \quad \pi_0 = 1. \tag{3.21}$$

(2) THE REPRESENTATION SATISFIES THE BIRTH AND DEATH EQUATIONS:

The function defined by the right-hand side of equation (3.14) is an analytic function of t, provided $Re(t) > 0$. Differentiating under the integral sign, we have

$$(3.22) \qquad \frac{\partial}{\partial t} P_{in}(t) = \pi_n \int_0^\infty e^{-xt} (-x) Q_i(x) Q_n(x) \, d\psi(x),$$

which becomes, upon using the recurrence relations (3.10) for $(-x)Q_n(x)$

$$\frac{\partial}{\partial t} P_{in}(t) = \pi_n \int_0^\infty e^{-xt} Q_i(x) \{\mu_n Q_{n-1}(x) - (\lambda_n + \mu_n) Q_n(x) + \lambda_n Q_{n+1}(x)\} d\psi(x),$$

$$= \frac{\pi_n}{\pi_{n-1}} \mu_n P_{i,n-1}(t) - (\lambda_n + \mu_n) P_{in}(t) + \frac{\pi_n}{\pi_{n+1}} \lambda_n P_{i,n+1}(t).$$

From equation (3.21), $\frac{\pi_n}{\pi_{n-1}} = \frac{\lambda_{n-1}}{\mu_n}$ and $\frac{\pi_n}{\pi_{n+1}} = \frac{\mu_{n+1}}{\lambda_n}$. Hence

$$\frac{\partial}{\partial t} P_{in}(t) = \lambda_{n-1} P_{i,n-1}(t) - (\lambda_n + \mu_n) P_{in}(t) + \mu_{n+1} P_{i,n+1}(t), \quad n \geq 1.$$

Similarly, using the recurrence relation for $(-x)Q_0(x)$, we find

$$\frac{\partial}{\partial t} P_{i0}(t) = \qquad -\lambda_0 P_{i0}(t) + \mu_1 P_{i1}(t).$$

Consequently the integral representation (3.14) satisfies the birth and death equations (3.1) and (3.2). The fact that the initial condition (3.3) is satisfied follows from the orthogonality condition (3.20).

If we substitute for $(-x)Q_i(x)$ instead of $(-x)Q_n(x)$ in equation (3.22) we obtain

(3.23) $$\frac{\partial}{\partial t} P_{in}(t) = \mu_i P_{i-1,n}(t) - (\lambda_i+\mu_i) P_{in}(t) + \lambda_i P_{i+1,n}(t), \quad i \geq 1,$$

(3.24) and $$\frac{\partial}{\partial t} P_{on}(t) = -\lambda_o P_{on}(t) + \lambda_o P_{1n}(t).$$

These equations are called the "backward equations" and may be written more compactly in the matrix form

(3.25) $$P'(t) = A\, P(t).$$

We have shown that the integral representation (3.14) also satisfies equation (3.25).

C. A NECESSARY AND SUFFICIENT CONDITION FOR UNIQUENESS

In this section we will prove the following:

(1) If equations (3.1), (3.2) and (3.3) have a unique solution, then the distribution function $\psi(x)$ is unique;

(2) if $\psi(x)$ is unique, then the series

$$\sum_{r=0}^{\infty} \left(\pi_r + \frac{1}{\lambda_r \pi_r} \right)$$

is divergent;

(3) if the above series is divergent then equations (3.1), (3.2) and (3.3) have a unique solution.

It follows that the three statements are equivalent. We shall use the so-called "contra-positive" argument in (2) and (3).

(1) UNIQUENESS OF THE $P_{in}(t)$ IMPLIES UNIQUENESS OF $\psi(x)$

Let us suppose that for two different distribution functions, $\psi_1(x)$ and $\psi_2(x)$,

$$\int_0^\infty e^{-xt} Q_i(x) Q_n(x)\, d\psi_1(x) = \int_0^\infty e^{-xt} Q_i(x) Q_n(x)\, d\psi_2(x),$$

for all i and n, and for all t in an interval $[a, b]$, where $0 \leq a < b$.

Then we have for all t in $[a, b]$

$$\int_0^\infty e^{-xt} Q_i(x) Q_n(x)\, d[\psi_1(x) - \psi_2(x)] = 0$$

and so by the uniqueness of the Laplace transform, (see, for example, Doetsch [13,VI,p 74]) $\psi_1(x) - \psi_2(x)$ can only be different from zero at the zeros of $Q_i(x) . Q_n(x)$. The moments of $\psi_1(x)$ and $\psi_2(x)$, obtained recursively from the $Q_n(x)$ (see section B, part 1, of this chapter,p79) are respectively equal and so all the moments of $\psi_1(x) - \psi_2(x)$ are zero. Hence $\psi_1(x) = \psi_2(x)$.

We have proved that if the solution to equations (3.1), (3.2) and (3.3) is unique then $\psi(x)$ is unique. Let us now find a necessary condition for $\psi(x)$ to be unique.

(2) A NECESSARY CONDITION FOR THE UNIQUENESS OF $\psi(x)$

At this point, we require some results which are proved by Shohat and Tamarkin [61,chapter2]. The results are:

(a) Let us define a function $\psi_n(x)$ to be a "distribution of order $n+1$ " associated with the Stieltjes moment problem if it consists of $n+1$ masses located on the real axis and has the correct moments c_k for $k=0,1,\cdots,2n$. Such a distribution can be found which has masses at the zeros of $Q_{n+1}(x)$.

(b) The sequence $\{\psi_n(x)\}$ of distribution functions of order $n+1$ which all have a mass at a fixed point x_0 , converges, as $n\to\infty$, to the solution $\psi_{min}(x)$, which has the maximum possible mass at x_0 .

(c) The largest mass which can be concentrated at a given point x_0 by any solution of the Stieltjes moment problem is $\rho(x_0)$, where

$$\rho(x) = \left\{\sum_{r=0}^{\infty} \pi_r Q_r(x)\right\}^{-1},$$

and there always exists a solution, $\psi_{max}(x)$, with the mass $\rho(x_0)$ concentrated at x_0 .

Let us define

$$\beta_r = \int_0^\infty \frac{1-Q_r(x)}{x}\, d\psi_n(x), \quad r=0,1,2,\ldots \tag{3.26}$$

Upon dividing equation (3.10) by x and integtating with respect to $\psi_n(x)$ we have

$$\int_0^\infty Q_r(x)\, d\psi_n(x) = \mu_r\int_0^\infty \frac{1-Q_{r-1}(x)}{x}\, d\psi_n(x) - (\lambda_r+\mu_r)\int_0^\infty \frac{1-Q_r(x)}{x}\, d\psi_n(x) + \lambda_r \int_0^\infty \frac{1-Q_{r+1}(x)}{x}\, d\psi_n(x). \tag{3.27}$$

When $r = n+1$, the left-hand side of equation (3.27) vanishes, because of the definition of $\psi_n(x)$. Hence we find, using equation (3.26),

$$\mu_{n+1}\beta_n - (\lambda_{n+1}+\mu_{n+1})\beta_{n+1} + \lambda_{n+1}\beta_{n+2} = 0, \quad n \geq 0.$$

It follows also that $\beta_0 = 0$ and $\lambda_0\beta_1 = 1$. Solving recursively for the β_r we obtain

$$\beta_r = \sum_{k=0}^{r-1} \frac{1}{\lambda_r \pi_r}$$

whence, from equation (3.26)

$$\int_0^\infty \frac{1 - Q_{n+1}(x)}{x} \, d\psi_n(x) = \sum_{r=0}^{n} \frac{1}{\lambda_r \pi_r}$$

Now $\int_0^\infty \frac{Q_{n+1}(x)}{x} d\psi_n(x) = 0$, and from result (a),

$$\psi_n(x) \to \psi_{min}(x) \text{ as } n \to \infty.$$

Thus

$$\int_0^\infty \frac{d\,\psi_{min}(x)}{x} = \sum_{r=0}^{\infty} \frac{1}{\lambda_r \pi_r}.$$

Consequently if $\sum_{r=0}^{\infty} \frac{1}{\lambda_r \pi_r} < \infty$ then $\psi_{min}(x)$ has no mass concentrated at $x = 0$.

However by result (c), there always exists a solution $\psi_{max}(x)$ which has mass

$$\rho = \left(\sum_{r=0}^{\infty} \pi_r\right)^{-1}$$

concentrated at $x = 0$. Hence if

$$\sum_{r=0}^{\infty} \left(\pi_r + \frac{1}{\lambda_r \pi_r}\right) < \infty$$

there are two distinct distributions, $\psi_{min}(x)$ and $\psi_{max}(x)$, with respect to which the system $\{Q_n(x)\}$ is orthogonal, the first having no mass at $x = 0$, the second having mass $\left(\sum_{r=0}^{\infty} \pi_r\right)^{-1}$ at $x = 0$. Taking the contrapositive statement, it follows that if $\psi(x)$ is unique, then

$$\sum_{r=0}^{\infty} \left(\pi_r + \frac{1}{\lambda_r \pi_r}\right) = \infty .$$

(3) A SUFFICIENT CONDITION FOR THE UNIQUENESS OF THE $P_{in}(t)$

Suppose we first consider the polynomials $Q_n(x)$. Let ξ_r, $r = 1, 2, \ldots, n$, be the zeros of $Q_n(x)$. Since the system $\{Q_n(x)\}$ is orthogonal on $[0, \infty)$ the ξ_r are all real and positive (see Szego [69,p43]). Since $Q_n(0) = 1$

$$Q_n(x) = \prod_{r=1}^{n} \left(1 - \frac{x}{\xi_r}\right).$$

(3.28) Hence $|Q_n(x)| = \prod_{r=1}^{n} \left|1 - \frac{x}{\xi_r}\right| \leq \prod_{r=1}^{n} \left(1 + \frac{|x|}{\xi_r}\right) = Q_n(-|x|).$

Upon putting $a_r = 1 + \frac{|x|}{\xi_r}$ in the inequality

$$\left(\prod_{r=1}^{n} a_r\right)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{r=1}^{n} a_r ,$$

we get

$$\prod_{r=1}^{n}\left(1+\frac{|x|}{\xi_r}\right) \leq \left\{\frac{1}{n}\left(n+|x|\sum_{r=1}^{n}\frac{1}{\xi_r}\right)\right\}^n.$$

Hence using the result (3.28),

$$(3.29) \quad Q_n(-|x|) \leq \left\{1+\frac{|x|}{n}\sum_{r=1}^{n}\frac{1}{\xi_r}\right\}^n \leq e^{|x|\sum_{r=1}^{n}\frac{1}{\xi_r}}.$$

Now $Q_n'(0) = -\sum_{r=1}^{n}\frac{1}{\xi_r}$. Using equations (3.28) and (3.29), we have

$$|Q_n(x)| \leq e^{-|x|Q_n'(0)}.$$

Also if $x_0 > 0$, $Q_n(-x_0) = \prod_{r=1}^{n}\left(1+\frac{x_0}{\xi_r}\right) \geq 1 + x_0\sum_{r=1}^{n}\frac{1}{\xi_r}$.
We have shown that for $Q_n(x)$ to be bounded for some $x < 0$, it is necessary that $-Q_n'(0)$ be bounded, in which case $Q_n(x)$ is bounded for all x . Let us find an expression for $-Q_n'(0)$.

From equation (3.10) we have

$$-xQ_n(x)\pi_n = \lambda_n\pi_n\{Q_{n+1}(x)-Q_n(x)\} - \mu_n\pi_n\{Q_n(x)-Q_{n-1}(x)\},$$

(3.30) whence $-xQ_n(x)\pi_n = \lambda_n\pi_n\{Q_{n+1}(x)-Q_n(x)\} - \lambda_{n-1}\pi_{n-1}\{Q_n(x)-Q_{n-1}(x)\},$

since $\lambda_{n-1}\pi_{n-1} = \mu_n\pi_n$, from equation (3.21).
Similarly we find that

$$(3.31) \quad -xQ_0(x)\pi_0 = \lambda_0\pi_0\{Q_1(x)-Q_0(x)\}.$$

Summing equations (3.30) for $n = 1, 2, \dots, r$ and adding equation (3.31), we find

$$(3.32) \qquad -x \sum_{j=0}^{r} \pi_j Q_j(x) \quad = \quad \lambda_n \pi_n \{ Q_{n+1}(x) - Q_n(x) \} .$$

Differentiating equation (3.32) with respect to x, putting $x = 0$, and summing over values of r from 0 to n, we finally obtain

$$(3.33) \qquad \sum_{r=0}^{n} \frac{1}{\lambda_r \pi_r} \sum_{j=0}^{r} \pi_j \quad = \quad - Q'_{n+1}(0).$$

Because of the previous result, it follows that $Q_n(x)$ is bounded for some $x < 0$ only if

$$\sum_{r=0}^{\infty} \frac{1}{\lambda_r \pi_r} \sum_{j=0}^{\infty} \pi_j \quad < \quad \infty .$$

We will assume that the solution to equations (3.8) and (3.9) is not unique, and that $P(t)$ and $P^{(1)}(t)$ are two distinct matrixes which satisfy these equations. We define

$$(3.34) \qquad R_{in}(p) \quad = \quad \int_0^\infty e^{-pt} \{ P_{in}(t) - P^{(1)}_{in}(t) \} dt .$$

Then since $0 \leq \sum_{j=0}^{n} P_{ij}(t) \leq 1$, for all $t \geq 0$ and every i, n,

$$(3.35) \qquad \left| \sum_{j=0}^{n} R_{ij}(p) \right| \quad \leq \quad \frac{2}{p} .$$

Denote by $R(p)$ the infinite matrix $\{R_{in}(p)\}, i, n = 0, 1, \dots$ Taking Laplace transforms in equations (3.8) and (3.25) respectively, and using the initial condition (3.9) and the definition (3.34) we obtain

(3.36) $$p\,R(p) = R(p)\,A$$

(3.37) and $$p\,R(p) = A\,R(p).$$

The most general solution to equation (3.36) is

$$R_{in}(p) = \pi_n Q_n(-p) f_i(p),$$

and the most general solution to equation (3.37) is

$$R_{in}(p) = Q_i(-p) f_n(p),$$

where $f_i(p)$ and $f_n(p)$ are arbitrary functions. Thus the most general solution to both equations (3.36) and (3.37) is

(3.38) $$R_{in}(p) = \pi_n Q_i(-p) Q_n(-p) R_{oo}(p).$$

Substituting equation (3.38) in the inequality (3.35), we have

$$\left| \sum_{j=0}^{n} Q_i(-p) Q_j(-p) \pi_j R_{oo}(p) \right| \leq \frac{2}{p}.$$

Since $R_{oo}(p_o) \neq 0$ for some $p_o > 0$, the above inequality implies that

(a) $Q_i(-p_o)$ is bounded for all i,

(b) $\sum_{j=0}^{\infty} Q_j(-p_o)\,\pi_j$ converges.

The result (a) implies that

$$\sum_{i=0}^{\infty} \frac{1}{\lambda_i \pi_i} \sum_{j=0}^{i} \pi_j < \infty$$

using the previous result, whence

$$\sum_{r=0}^{\infty} \frac{1}{\lambda_r \pi_r} < \infty .$$

The result (b) implies that $\sum_{j=0}^{\infty} \pi_j < \infty$.

We have proved that if the solution to equations (3.1), (3.2) and (3.3) is not unique then

$$\sum_{r=0}^{\infty} \left(\frac{1}{\lambda_r \pi_r} + \pi_r \right) < \infty .$$

It follows that if this series diverges, the solution to the birth and death equations is unique. This completes our investigation of the existence and uniqueness of the integral representation for $P_{in}(t)$.

In order to find the $P_{in}(t)$ we must find the distribution function $\psi(x)$. For many functions λ_n and μ_n it turns out that the polynomials $Q_n(x)$ are classical orthogonal polynomials, and the distribution function with respect to which they are orthogonal is known. There are, however, many important problems for which the $Q_n(x)$ are not classical orthogonal polynomials, Accordingly Karlin and McGregor [32, section 8] have obtained an expression for the Stieltjes transform of $\psi(x)$, defined as

$$(3.39) \qquad B(s) = \int_0^{\infty} \frac{1}{x+s} \, d\psi(x) .$$

The expression for $B(s)$ is known provided the asymptotic behaviour of the $Q_n(x)$ and certain "associated" (that is, obeying the same recurrence relation with different initial conditions) polynomials $Q_n^{(0)}(x)$ as $n \to \infty$, for $-\alpha \le x \le -\beta$, where $0 \le \beta < \alpha$, is known.

By using the classical inversion formula for the Stieltjes transform we may then find $\psi(x)$. However the problem of finding these asymptotic forms from the recurrence relations does not appear to be easy, in general, so we will not pursue the matter further.

D. SOME EXAMPLES

1 THE QUEUE WITH INFINITELY MANY SERVERS

If we let the number of servers, m, be infinite, no queue will form, and the number in the system is equal to the number of customers being served. This problem was first considered as such by Feller [19], although Kendall [43] solved a problem in population growth which is very similar to a particular case of the queue with an infinite number of servers. Both obtained expressions for the $P_{in}(t)$ by using generating function techniques.

The function μ_n in this case is given by $\mu_n = n\mu$.

We shall consider two types of functions λ_n:

(a) $\lambda_n = \lambda$; (b) $\lambda_n = (n+1)\lambda$.

(a) STATE-INDEPENDENT ARRIVAL PROCESS

Case (a) corresponds to the queue in which the arrival process is state-independent, and this is the example which Feller considered. Case (b) is a queue in which customers are encouraged by the sight of a large number in the system. This phenomenon is often encountered at fair-ground sideshows. For each case the polynomials $Q_n(x)$ are classical.

In example (a) the $Q_n(x)$ are given by

$$Q_n(x) = c_n\left(\frac{x}{\mu}, \frac{\lambda}{\mu}\right),$$

where the c_n are the Poisson-Charlier polynomials. They satisfy the recurrence relations

$$c_0(x, a) = 1$$

$$\text{and} \quad -x\, c_n(x, a) = n\, c_{n-1}(x,a) - (n+a)\, c_n(x,a) + a\, c_{n+1}(x,a), \quad n \geq 0,$$

and are orthogonal with respect to the distribution which is a step-function with jumps

$$e^{-a}\frac{a^x}{x!} \quad \text{at} \quad x = 0, 1, 2, \cdots .$$

(see, for example, Erdelyi [18,Vol2,pp226-7]). In this case from equation (3.21) we have

$$\pi_n = \frac{\lambda^n}{n!\,\mu^n},$$

whence the series $\sum_{r=1}^{\infty} \frac{1}{\lambda_r \pi_r}$ is divergent. It follows that the $P_{in}(t)$ are unique. Using equation (3.14) we find

$$(3.40) \quad P_{in}(t) = \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!}\, e^{-\frac{\lambda}{\mu}} \sum_{r=0}^{\infty} \frac{\left(\frac{\lambda}{\mu}e^{-\mu t}\right)^r}{r!}\, c_i\left(r, \frac{\lambda}{\mu}\right) c_n\left(r, \frac{\lambda}{\mu}\right).$$

From the properties of the Poisson-Charlier polynomial we have

(3.41) $$c_n(r, a) = c_r(n, a)$$

(3.42) and $$\sum_{n=0}^{\infty} c_n(x, a)\frac{z^n}{n!} = e^z\left(1 - \frac{z}{a}\right)^x.$$

Putting $i = 0$ in equation (3.40) and using equations (3.41) and (3.42) we obtain

(3,43) $$P_{on}(t) = \frac{(\lambda/\mu)^n}{n!} \exp\left\{-\frac{\lambda}{\mu}(1-e^{-\mu t})\right\} \cdot (1-e^{-\mu t})^n.$$

(b) A QUEUE WITH ENCOURAGEMENT

Considering now example (b), we distinguish three cases:

(i) $\lambda < \mu$; (ii) $\lambda = \mu$; (iii) $\lambda > \mu$.

For cases (i) and (iii) the $Q_n(x)$ may be expressed in terms of the Meixner polynomials $m_n(x, \beta, \gamma)$ which, when $\beta = 1$, satisfy the recurrence relation:

$$-x\left(\frac{1-\gamma}{\gamma}\right)m_n(x, 1, \gamma) = \frac{1}{\gamma} n^2 m_{n-1}(x, 1, \gamma) - \left(\frac{n}{\gamma} + n + 1\right) m_n(x, 1, \gamma) + m_{n+1}(x, 1, \gamma),$$

with $m_{-1}(x, 1, \beta) = 0$. (see Meixner [52]).

The $m_n(x, \beta, \gamma)$, defined for $\beta > 0$ and $0 < \gamma < 1$, are orthogonal with respect to a step-function with jumps

$$(1-\gamma)^{-\beta} \frac{\Gamma(\beta + x)}{\Gamma(\beta)\, x!} \gamma^x \text{ at } x = 0, 1, 2, \cdots.$$

Furthermore the m_n may be shown to satisfy the identities

$$(3.44)\qquad \Gamma(\beta+x)\, m_n(x;\beta,\gamma) = \Gamma(\beta+n)\, m_x(n;\beta,\gamma)$$

$$(3.45)\qquad \text{and } \sum_{n=0}^{\infty} m_n(x;\beta,\gamma)\frac{z^n}{n!} = \left(1-\frac{z}{\gamma}\right)^x (1-z)^{-x-\beta},$$

(see, for example Erdelyi [18,Vol2,p225]).

Let us work out the $P_{in}(t)$ for case (i). Since, from equation (3.21), $\pi_n = \left(\frac{\lambda}{\mu}\right)^n$, the series $\sum_{n=0}^{\infty}\left(\pi_n + \frac{1}{\lambda_n \pi_n}\right)$ is divergent for all three cases. It is easily verified that the polynomials

$$Q_n(x) = \frac{1}{n!}\, m_n\left(\frac{x}{\mu-\lambda};1,\frac{\lambda}{\mu}\right)$$

satisfy the recurrence relations

$$Q_0(x) = 1$$

$$-x\,Q_n(x) = n\mu\, Q_{n-1}(x) - (n\mu+n\lambda+\lambda)\,Q_n(x) + (n+1)\lambda\, Q_{n+1}(x),$$
$$n \geq 0.$$

Using equation (3.14), the $P_{in}(t)$ are given by

$$(3.46)\qquad P_{in}(t) = \left(\frac{\lambda}{\mu}\right)^n\left(1-\frac{\lambda}{\mu}\right)^{-1}\sum_{r=0}^{\infty}\left(\frac{\lambda}{\mu}e^{-(\mu-\lambda)t}\right)^r m_i\left(r;1,\frac{\lambda}{\mu}\right) m_n\left(r,1,\frac{\lambda}{\mu}\right).$$

Now for the case $i = 0$, if we use the relations (3.44), and (3.45) with $\beta = 1$, then equation (3.46) becomes

$$(3.47)\qquad P_{0n}(t) = \left(\frac{\lambda}{\mu}\right)^n\left(1-\frac{\lambda}{\mu}\right)^{-1}\left(1-e^{-(\mu-\lambda)t}\right)^n\left(1-\frac{\lambda}{\mu}e^{-(\mu-\lambda)t}\right)^{-n-1}.$$

For case (ii) the $Q_n(x)$ are the well-known Laguerre polynomials $L_n(x)$, defined by

$$L_n(x) = \frac{e^x}{n!}\frac{d^n}{dx^n}(e^{-x}x^n).$$

These polynomials are orthogonal with respect to a distribution function which has the continuous derivative

$$\psi'(x) = e^{-x} \quad \text{on} \quad 0 \leq x < \infty$$

(see Erdelyi [18,Vol2,p188et seq]).

Since the polynomials $L_n\left(\frac{x}{\lambda}\right)$ satisfy the recurrence relations

$$L_0\left(\frac{x}{\lambda}\right) \equiv 1$$

$$-x\, L_n\left(\frac{x}{\lambda}\right) = n\lambda L_{n-1}\left(\frac{x}{\lambda}\right) - \{n\lambda + (n+1)\lambda\} L_n\left(\frac{x}{\lambda}\right) + (n+1)\lambda L_{n-1}\left(\frac{x}{\lambda}\right), \quad n \geq 0,$$

the $P_{in}(t)$ are given by

(3.48) $$P_{in}(t) = \int_0^\infty e^{-x(t+1)} L_i\left(\frac{x}{\lambda}\right) L_n\left(\frac{x}{\lambda}\right) dx.$$

In particular

$$P_{0n}(t) = \int_0^\infty e^{-x(t+1)} L_n\left(\frac{x}{\lambda}\right) dx.$$

Using Erdelyi [17,Vol 1,p175] to evaluate the Laplace transform of $e^{-x} L_n\left(\frac{x}{\lambda}\right)$, we find

(3.49) $$P_{0n}(t) = \frac{(\lambda t)^n}{(1+\lambda t)^{n+1}}.$$

Both examples (a) and (b) may be solved using generating function techniques. Let

$$(3.50) \qquad \lambda_n = an + b, \qquad \mu_n = \mu n .$$

Then equations (3.1) and (3.2) become

$$(3.51) \quad \frac{\partial}{\partial t} P_{in}(t) = \{(n-1)a + b\} P_{i,n-1}(t) - (an + b + \mu n) P_{in}(t) + (n+1)\mu P_{i,n+1}(t), \quad n \geq 1,$$

$$(3.52) \text{ and } \frac{\partial}{\partial t} P_{i0}(t) = \qquad - b\, P_{i0}(t) + \mu P_{i1}(t).$$

Multiplying equation (3.51) by z^n, summing over all positive integral values of n, and adding equation (3.52) we find

$$(3.53) \qquad \frac{\partial}{\partial t} \Pi(z,t) = \{az^2 - (a+\mu)z + \mu\} \frac{\partial}{\partial t} \Pi(z,t) + b(z-1)\Pi(z,t).$$

Now the partial differential equation (3.53) has associated Lagrange equations

$$(3.54) \qquad \frac{dt}{1} = \frac{dz}{-(1-z)(\mu - az)} = \frac{d\Pi}{-b(1-z)\Pi}$$

(see, for example, Nielson [56,p191]). If we put $a=0$, $b=\lambda$, we have example (a) which we considered before. In this case equations (3.54) have the two independent solutions

$$u_1 = \frac{1}{\mu}(1-z)e^{-\mu t} \quad \text{and} \quad u_2 = \Pi e^{-\frac{\lambda}{\mu} z}.$$

It follows that the general solution to equation (3.53) is

$$\Pi(z,t) = e^{\frac{\lambda}{\mu}z}\,\phi([1-z]e^{-\mu t}),$$

where ϕ is an arbitrary function. From the initial condition (3.3):

$$\Pi(z,0) = z^{i},$$

whence $\phi(w) = e^{-\frac{\lambda}{\mu}(1-w)}(1-w)^{i}.$

Consequently the generating function $\Pi(z,t)$ is

(3.55) $$\Pi(z,t) = e^{\frac{\lambda}{\mu}z}\exp\left[-\frac{\lambda}{\mu}\{1-(1-z)e^{-\mu t}\}\right].\{1-(1-z)e^{-\mu t}\}^{i}.$$

The $P_{in}(t)$ are now found by picking out coefficients of z^{n} in equation (3.55). In particular, when $i = 0$

$$P_{on}(t) = \frac{1}{n!}\left\{\frac{\lambda}{\mu}(1-e^{-\mu t})\right\}^{n}\exp\left\{-\frac{\lambda}{\mu}(1-e^{-\mu t})\right\},$$

which agrees with equation (3.43).

Substituting $a = b = \lambda$ in equation (3.50) we get the second example (b). The two independent solutions to equations (3.54) are now, for $\lambda < \mu$,

$$u_1 = \frac{1-z}{\mu-\lambda z}e^{-(\mu-\lambda)t} \quad \text{and} \quad u_2 = (\mu-\lambda z)\Pi.$$

Proceeding as before, we obtain

$$\Pi(z,t) = \left(1-\frac{\lambda}{\mu}z\right)^{-1}\left\{1-\frac{\lambda}{\mu}\left[\frac{\mu-\lambda z-\mu(1-z)e^{-(\mu-\lambda)t}}{\mu-\lambda z-\lambda(1-z)e^{-(\mu-\lambda)t}}\right]\right\}\left\{\frac{\mu-\lambda z-\mu(1-z)e^{-(\mu-\lambda)t}}{\mu-\lambda z-\lambda(1-z)e^{-(\mu-\lambda)t}}\right\}^{i}.$$

Simplifying the above, we find that

$$(3.56)\quad \Pi(z,t) = \left(1-\frac{\lambda}{\mu}\right)\left(1-\frac{\lambda}{\mu}e^{-(\mu-\lambda)t}\right)^{-1}\left(1-\frac{\lambda(1-e^{-(\mu-\lambda)t})}{\mu-\lambda e^{-(\mu-\lambda)t}}z\right)^{-1}\left\{\frac{\mu-\lambda z-\mu(1-z)e^{-(\mu-\lambda)t}}{\mu-\lambda z-\lambda(1-z)e^{-(\mu-\lambda)t}}\right\}^{i}$$

Thus, for $i = 0$,

$$P_{on}(t) = \left(\frac{\lambda}{\mu}\right)^{n}\left(1-\frac{\lambda}{\mu}\right)\left(1-\frac{\lambda}{\mu}e^{-(\mu-\lambda)t}\right)^{n-1}\left(1-e^{-(\mu-\lambda)t}\right)^{n}.$$

As for example (a), this agrees with the result which we obtained from the integral representation, given by equation (3.47).

Finally, let us solve for the case $\lambda = \mu$. The solutions to the Lagrange equations are

$$u_1 = t + \frac{1}{\lambda(1-z)}, \qquad u_2 = (1-z).$$

(3.57) We find $\Pi(z,t) = (1+\lambda t(1-z))^{-1}\left(\dfrac{\lambda t+z(1-\lambda t)}{1+\lambda t(1-z)}\right)^{i}$,

whence $P_{on}(t) = (1+\lambda t)^{-1}\left(\dfrac{\lambda t}{1+\lambda t}\right)^{n}$, which is identical to equation (3.49).

2. THE QUEUE WITH A FINITE NUMBER OF SERVERS

A further example which we will consider is the many-server queue with a state-independent arrival process. We have

$$\lambda_n = \lambda \quad \text{for all } n \tag{3.58}$$
$$\mu_n = \begin{cases} n\mu & \text{if } n \le m, \\ m\mu & \text{if } n > m. \end{cases}$$

Many server queues have been investigated by Kendall [44] and by Kiefer and Wolfowitz [47]. However the first explicit solution for the particular queue under consideration was obtained by Karlin and McGregor [33]. Saaty [59] found expressions for the Laplace transforms of the $P_{in}(t)$ using generating functions. Keilson [38] also investigated this problem.

The polynomials $Q_n(x)$ are found to be classical in the sense that they are expressible in terms of the well-known Chebyshev polynomials. In order to find the distribution $\psi(x)$ we will make use of a very important and useful result of Karlin and McGregor. As we will not have space to repeat the proof which is lengthy, we will simply enunciate the theorem.

<u>THEOREM 3.1</u> : Let $A^{(r)}$ be the matrix obtained from the matrix A, defined in equation (3.8), by omitting the first r rows and columns. Let $\{Q_n^{(r)}(x)\}$ be the system of polynomials defined by the recurrence relations

$$-x\,Q^{(r)}(x) = A^{(r)}Q^{(r)}(x)\,; \quad Q_0^{(r)}(x) \equiv 1.$$

(1) The system $\{Q_n^{(r)}(x)\}$ is orthogonal with respect

to a unique distribution $\psi^{(r)}(x)$, defined on the interval $[0, \infty)$, and not supported by a finite set of points, if and only if

$$\sum_{r=0}^{\infty} \left(\pi_r + \frac{1}{\lambda_r \pi_r} \right) = \infty .$$

(2) If $B^{(r)}(s)$ denotes the Stieltjes transform of $\psi^{(r)}(x)$ (see equation (3.39)), then

(3.59) $$\{ B^{(r)}(s) \}^{-1} = \lambda_r + \mu_r + s - \lambda_r \mu_{r+1} B^{(r+1)}(s).$$

The proof follows from the fact that the process is a stationary Markov process. Since, for the m-server queue, in view of equations (3.59), the matrixes $A^{(r)}, r \geq m$, are identical, we have

$$B^{(m)}(x) = B^{(m+1)}(x).$$

Putting this in equation (3.59) we get

$$\lambda m \mu [B^m(s)]^2 - (\lambda + m\mu + s) B^m(s) + 1 = 0,$$

(3.60) whence $$B^{(m)}(s) = \frac{1}{2\lambda m \mu} \left\{ \lambda + m\mu + s - \sqrt{(\lambda + m\mu + s)^2 - 4\lambda m \mu} \right\}.$$

We have taken the negative square root because, using the definition (3.39), we must have

$$B^{(m)}(s) \to 0 \quad \text{as } s \to \infty .$$

We find $B(s)$, the Stieltjes transform of $\psi(x)$ (defined by equation (3.39)), by successive application of formula (3.59). To illustrate the method, let us solve for the case $m = 1$.

From equation (3.60) we have, putting $m = 1$

$$B^{(1)}(s) = \frac{1}{2\lambda\mu}\left\{\lambda + \mu + s - \sqrt{(\lambda+\mu+s)^2 - 4\lambda\mu}\right\}.$$

Using equation (3.59), we find

$$B(s) = B^{(0)}(s) = \frac{1}{2\mu s}\left\{-\lambda + \mu - s + \sqrt{(\lambda+\mu+s)^2 - 4\lambda\mu}\right\}.$$

Now $B(s)$ has a simple pole at $s = 0$, where the residue is

$$\rho = \begin{cases} 0 & \text{if } \lambda \geq \mu, \\ 1 - \frac{\lambda}{\mu} & \text{if } \lambda < \mu. \end{cases}$$

Consequently $\psi(x)$ has a jump of $1 - \frac{\lambda}{\mu}$ at $x = 0$ if $\lambda < \mu$, but no jump at $x = 0$ if $\lambda \geq \mu$. Let $\psi^*(x)$ be the distribution with this jump removed, and $B^*(s)$ the corresponding Stieltjes transform. Since $\psi^*(x)$ is absolutely continuous with density function $\psi_1^*(x)$, say, we may use the Stieltjes inversion formula for absolutely continuous functions:

$$(3.61) \qquad \psi_1^*(x) = \lim_{y \to 0+} \frac{B^*(-x-iy) - B^*(-x+iy)}{2\pi i}.$$

(see, for example, Widder [66,p340]). Removing the contribution to $B(s)$ from the jump at $x = 0$, we get

$$(3.62) \qquad \beta(s) = \frac{\sqrt{(\lambda+\mu+s)^2 - 4\lambda\mu}}{2\mu s} .$$

Now applying formula (3.62) to the function $\beta^*(s)$, given in equation (3.62), we obtain

$$(3.63) \qquad \psi_1^*(x) = \begin{cases} \frac{1}{2\pi\mu x}\sqrt{4\lambda\mu-(\lambda+\mu-x)^2} & \text{if } 4\lambda\mu \geq (\lambda+\mu-x)^2, \\ 0 & \text{otherwise.} \end{cases}$$

Since from equation (3.21),

$$\pi_n = \left(\frac{\lambda}{\mu}\right)^n,$$

we finally obtain, after using equations (3.14), (3.63) and (3.19)

$$(3.64) \qquad P_{in}(t) = \left(\frac{\lambda}{\mu}\right)^n \int_{(\sqrt{\lambda}-\sqrt{\mu})^2}^{(\sqrt{\lambda}+\sqrt{\mu})^2} e^{-xt} Q_i(x) Q_n(x) \frac{\sqrt{4\lambda\mu-(\lambda+\mu-x)^2}}{2\pi\mu x}\, dx + \\ + \left(\frac{\lambda}{\mu}\right)^n \left(1-\frac{\lambda}{\mu}\right) S\left(\frac{\mu}{\lambda}\right),$$

where $S(x) = \begin{cases} 0 & \text{if } x \leq 1 \\ 1 & \text{if } x > 1 \end{cases}$, and the $Q_n(x)$ are given by

$$-x Q_n(x) = \mu Q_{n-1}(x) - (\lambda+\mu) Q_n(x) + \lambda Q_{n+1}(x),$$

$$\text{with} \quad Q_0(x) \equiv 1, \quad Q_{-1}(x) \equiv 0 .$$

<u>3. THE BUSY PERIOD FOR THE MANY-SERVER QUEUE</u>

In view of the relationship between equations (3.1), (3.2) and (3.4), (3.5), the relation (3.60) also enables us to find the distribution $\hat{\psi}(x)$ for the polynomials $\hat{Q}_n(x)$ for the modified process. We are thus able to find the probability distribution function of the busy periods for all queueing processes specified by equations (3.1), (3.2) and (3.3), provided the function $\psi(x)$ is known. In the case of many servers, we must define exactly what is meant by a busy period, that is, whether all servers are busy, or only a certain number. To illustrate the method, let us consider the two-server queue.

Let us define a busy period ω to be the length of a period at least one server is busy. Then the modified process commences in the state $n = 1$. The probability distribution function of ω is, from equation (1.37)

$$F_\omega(t) = \hat{P}_{10}(t).$$

Using equation (3.6), this becomes

(3.65) $$F_\omega(t) = \mu \int_0^t \hat{P}_{11}(t')\,dt'.$$

From part (1) of Theorem 3.1, the $\hat{P}_{in}(t)$ for the modified process are expressible in the integral representation (3.14). Consequently we may write equation (3.65) as

$$F_\omega(t) = \mu \int_0^t \int_0^\infty e^{-xt'}\,d\hat{\psi}(x)\,dt'.$$

Interchanging the order of integration we have

$$(3.66) \qquad F_\omega(t) = \mu \int_0^\infty \frac{1-e^{-xt}}{x}\, d\hat{\psi}(x).$$

For the two server process, from equation (3.60)

$$B^{(2)}(s) = \frac{1}{4\lambda\mu}\{\lambda + 2\mu + s - \sqrt{(\lambda+2\mu+s)^2 - 8\lambda\mu},$$

whence, using equation (3.59)

$$\hat{B}(s) = B^{(1)}(s) = \frac{\lambda + s - \sqrt{(\lambda+2\mu+s)^2 - 8\lambda\mu}}{2\mu(\lambda-\mu-s)}.$$

Since $\hat{B}(s)$ has a simple pole at $s = \lambda - \mu$, where the residue is $1-\frac{2\lambda}{\mu}$ if $2\lambda < \mu$ but zero otherwise, $\hat{\psi}(x)$ has a jump of magnitude $1-\frac{2\lambda}{\mu}$ at $x = \mu - \lambda$ if $2\lambda < \mu$, but no jump otherwise. Removing this jump from $\hat{\psi}(x)$, and using formula (3.61) to find the remaining part $\hat{\psi}^*(x)$, we find that

$$\hat{\psi}_1^*(x) = \begin{cases} \dfrac{\sqrt{8\lambda\mu - (\lambda+2\mu-x)^2}}{2\pi\mu(x-\mu+\lambda)} & \text{if } 8\lambda\mu \geq (\lambda+2\mu-x)^2; \\ 0 & \text{otherwise.} \end{cases}$$

Thus from equation (3.66), ω has the probability distribution function

$$(3.67) \qquad F_\omega(t) = \int_{(\sqrt{\lambda}-\sqrt{\mu})^2}^{(\sqrt{\lambda}+\sqrt{\mu})^2} (1-e^{-xt}) \frac{\sqrt{8\lambda\mu - (\lambda+2\mu-x)^2}}{2\pi x(x-\mu+\lambda)}\, dx + \\ + (1-e^{-(\mu-\lambda)t}) \frac{\mu-2\lambda}{\mu-\lambda}\, \delta\!\left(\frac{\mu}{2\lambda}\right).$$

If we now let $t \to \infty$ we find that the distribution is "honest" (see chapter 1,p35) if $2\lambda \leq \mu$. A similar argument may be used to find the probability distribution function of the length of the time when both servers are busy.

4. A QUEUE WITH A SPECIAL ARRIVAL PROCESS.

The above examples of queues with a finite or infinite number of servers have been studied extensively by Karlin and McGregor [33] and [34]. Let us consider now a queue which has not been studied at all in the literature to date (although Karlin and McGregor [36] have considered a similar problem in the theory of random walks) but to which the results and methods of Karlin and McGregor are applicable. We put

$$(3.68) \quad \mu_n = \mu, \text{ for all } n, \text{ and } \lambda_n = \begin{cases} \lambda_1 & \text{if } n \text{ is odd} \\ \lambda_2 & \text{if } n \text{ is even.} \end{cases}$$

It is clear that for this process the matrices $A^{(r)}$ and $A^{(r+2)}$ (see Theorem 3.1) are identical. It follows from equation (3.59) that

$$(3.69) \quad \frac{1}{B(s)} = \frac{1}{B^{(2)}(s)} = \lambda_2 + s - \lambda_2 \mu B^{(1)}(s),$$

$$(3.70) \quad \frac{1}{B^{(1)}(s)} = \lambda_1 + \mu + s - \lambda_1 \mu B^{(2)}(s),$$

$$(3.71) \text{ and } \frac{1}{B^{(2)}(s)} = \lambda_2 + \mu + s - \lambda_2 \mu B^{(3)}(s).$$

(3.72) Also $B^{(1)}(s) = B^{(3)}(s)$.

Solving equations (3.69) - (3.72) for $B(s)$ we find that $B(s)$ satisfies the quadratic equation

$$\mu s(\lambda_2+\mu+s)[B(s)]^2+[(\lambda_1+s)(\lambda_2+s)-\mu^2]B(s)-(\lambda_1+\mu+s) = 0.$$

(3.74) Thus $B(s) = \dfrac{\mu^2-(\lambda_1+s)(\lambda_2+s)-\sqrt{[(\lambda_1+s)(\lambda_2+s)-\mu^2]^2+4\mu s(\lambda_1+\mu+s)(\lambda_2+\mu+s)}}{2\mu s(\lambda_2+\mu+s)}$.

$B(s)$ has simple poles at $s=0$ and $s=-(\lambda_2+\mu)$. Considering the pole at $s=0$, the residue is

$$\rho(0) = \begin{cases} \dfrac{\mu^2-\lambda_1\lambda_2}{\mu(\lambda_2+\mu)} & \text{if } \lambda_1\lambda_2<\mu^2, \\ 0 & \text{if } \lambda_1\lambda_2 \geq \mu^2. \end{cases}$$

For the pole at $s=-(\lambda_2+\mu)$, the residue is

$$\rho(-\lambda_2-\mu) = \begin{cases} \dfrac{\lambda_2-\lambda_1}{\lambda_2+\mu} & \text{if } \lambda_1<\lambda_2 \\ 0 & \text{if } \lambda_1 \geq \lambda_2 . \end{cases}$$

Consequently $\psi(x)$ has a jump of magnitude $\dfrac{\mu^2-\lambda_1\lambda_2}{\mu(\lambda_2+\mu)}$ at $x=0$ if $\lambda_1\lambda_2<\mu^2$, but no jump otherwise, and a jump of magnitude $\dfrac{\lambda_2-\lambda_1}{\lambda_2+\mu}$ at $x=\lambda_2+\mu$ if $\lambda_1<\lambda_2$, but no jump otherwise. Removing these jumps from $\psi(x)$ we obtain an absolutely continuous function $\psi^*(x)$, to which we may apply formula (3.61). We deduce

(3.74) $$\psi_1^*(x) \doteq \begin{cases} \dfrac{\sqrt{4\mu x(\lambda_1+\mu-x)(\lambda_2+\mu-x)-\{(\lambda_1-x)(\lambda_2-x)-\mu^2\}^2}}{2\pi\mu x(\lambda_2+\mu-x)} & \text{whenever the expression inside the square root is positive,} \\ 0 & \text{otherwise.} \end{cases}$$

Now in this case we have from equation (3.21)

$$\pi_n = \begin{cases} \left(\frac{\sqrt{\lambda_1\lambda_2}}{\mu}\right)^n & \text{if } n \text{ is even,} \\ \left(\frac{\sqrt{\lambda_1\lambda_2}}{\mu}\right)^n \frac{\lambda_1}{\mu} & \text{if } n \text{ is odd.} \end{cases}$$

Thus the series $\sum_{n=0}^{\infty}\left(\pi_n + \frac{1}{\lambda_n \pi_n}\right)$ is always divergent. The $P_{in}(t)$ are unique, and from equations (3.10), (3.14), (3.74) and (3.19), are given, for even values of n, by

$$\begin{aligned} P_{in}(t) = \left(\frac{\sqrt{\lambda_1\lambda_2}}{\mu}\right)^n \int_I e^{-xt} Q_i(x) Q_n(x) \frac{\sqrt{4\mu x(\lambda_1+\mu-x)(\lambda_2+\mu-x) - \{(\lambda_1-x)(\lambda_2-x)-\mu^2\}^2}}{2\pi\mu x(\lambda_2+\mu-x)}\,dx \\ + \left(\frac{\sqrt{\lambda_1\lambda_2}}{\mu}\right)^n \frac{\mu^2-\lambda_1\lambda_2}{\mu(\lambda_2+\mu)} S\left(\frac{\mu^2}{\lambda_1\lambda_2}\right) + \left(\frac{\sqrt{\lambda_1\lambda_2}}{\mu}\right)^n Q_i(\lambda_2+\mu) Q_n(\lambda_2+\mu) \frac{\lambda_2-\lambda_1}{\lambda_2+\mu} S\left(\frac{\lambda_2}{\lambda_1}\right). \end{aligned} \tag{3.75}$$

where I is the interval or intervals, on the positive x-axis, where $4\mu x(\lambda_1+\mu-x)(\lambda_2+\mu-x) \geq \{(\lambda_1-x)(\lambda_2-x)-\mu^2\}^2$, and $Q_n(x)$ are the polynomials given by the recurrence relation:

$$\begin{aligned} -x\,Q_n(x) &= \mu Q_{n-1}(x) - (\lambda_1+\mu)Q_n(x) + \lambda_1 Q_{n+1}(x) \quad \text{for } n \text{ odd,} \\ -x\,Q_n(x) &= \mu Q_{n-1}(x) - (\lambda_2+\mu)Q_n(x) + \lambda_2 Q_{n+1}(x) \quad \text{for } n \text{ even,} \end{aligned}$$

with $Q_0(x) \equiv 1$ and $Q_{-1}(x) \equiv 0$.

For odd values of n, we must replace $\left(\frac{\sqrt{\lambda_1\lambda_2}}{\mu}\right)^n$ in equation (3.75) by $\left(\frac{\sqrt{\lambda_1\lambda_2}}{\mu}\right)^n \frac{\lambda_1}{\mu}$.

We may find the probability distribution function of the busy periods by similar reasoning. For the modified process, the distribution $\hat{\psi}(x)$ is identical with the function $\psi^{(1)}(x)$, defined in Theorem 3.1. Hence from equations (3.70), (3.71) and (3.72) we get

$$\hat{B}(s) = B^{(1)}(s) = \frac{(\lambda_1+\mu+s)(\lambda_2+\mu+s)+\mu(\lambda_2-\lambda_1) - \sqrt{\{(\lambda_1+\mu+s)(\lambda_2+\mu+s)+\mu(\lambda_2-\lambda_1)\}^2 - 4\lambda_2\mu(\lambda_1+\mu+s)(\lambda_2+\mu+s)}}{2\lambda_2\mu(\lambda_1+\mu+s)}.$$

Proceeding as before, we find that $\hat{\psi}^*(x)$ has continuous density

$$\hat{\psi}_1^*(x) = \frac{\sqrt{4\lambda_2\mu(\lambda_1+\mu-x)(\lambda_2+\mu-x) - \{(\lambda_1+\mu-x)(\lambda_2+\mu-x)+\mu(\lambda_2-\lambda_1)\}^2}}{2\pi\lambda_2\,\mu(\lambda_1+\mu-x)}$$

on the region $I : 4\lambda_2\mu(\lambda_1+\mu-x)(\lambda_2+\mu-x) \geq \{(\lambda_1+\mu-x)(\lambda_2+\mu-x)+\mu(\lambda_2-\lambda_1)\}^2$ together with a jump of magnitude $\frac{\lambda_2-\lambda_1}{2\lambda_2}\,\delta\left(\frac{\lambda_2}{\lambda_1}\right)$ at $x = \lambda_1+\mu$.
Thus from equation (3.66), which applies equally well to the queue we are considering, the probability distribution function of the busy periods is

$$F_w(t) = \int_I (1-e^{-xt})\,\frac{\sqrt{4\lambda_2\mu(\lambda_1+\mu-x)(\lambda_2+\mu-x)-\{(\lambda_1+\mu-x)(\lambda_2+\mu-x)+\mu(\lambda_2-\lambda_1)\}^2}}{2\pi\lambda_2 x(\lambda_1+\mu-x)}\,dx$$
$$+ (1-e^{-(\lambda_1+\mu)t})\frac{\lambda_2-\lambda_1}{2\lambda_2(\lambda_1+\mu)}\,\delta\left(\frac{\lambda_2}{\lambda_1}\right). \qquad (3.76)$$

This completes our investigation of queues with state-dependent arrival processes. In the next chapter we will consider what happens to the $P_{in}(t)$ as $t \to \infty$.

CHAPTER 4

THE ASYMPTOTIC BEHAVIOUR OF THE QUEUE

The behaviour of queues as $t \to \infty$, first considered in a special case by Erlang [16] (see Jensen [29] for a review of Erlang's works), has since been considered by many authors. Pollaczek [58] and Khinchine [46] considered the queue $M/G/1$, and independently derived a formula for the Laplace transform of the distribution function of a customer's waiting time. Kendall, [42] and [45], also considered this queue, and was the first to observe the fact that a Markov chain results by considering the number in the system at the instants at which customers depart. Lindley [37] derived an integral equation for the probability density function of the waiting time distribution. Takacs [62] obtained many results using combinatorial methods, as did Finch (see, for example [21]), and Morse [55] has considered the asymptotic behaviour of many different queueing processes. Important contributions concerning the limiting behaviour of queues have also come from Conolly [10], Jaiswal [28], Miller [53], Keilson [39] and Bhat [6].

Generally the limiting behaviour of the queues which we have discussed is simpler to determine than that of the queue $M/G/1$. For the queue with single arrivals and bulk service (chapter 2, section C) Bailey [3] derived the queue length distribution, while Downton [15] obtained the waiting time distribution. Foster [22] obtained these distributions for the queue with batch arrivals and single departures (chapter 2, section B). In the case when the arrival and departure processes are simple and state-dependent (chapter 3) the problem is trivial.

In this chapter we will find the limiting distribution of the queue length and the customer's waiting time for the queues which we have already discussed in the finite-time case. We will conclude with an example which is new.

A. EXISTENCE OF A LIMITING DISTRIBUTION FOR THE $P_{in}(t)$

In order to prove the existence of a limiting distribution for the $P_{in}(t)$, we use a theorem of P. Levy, which states:

THEOREM 41 : If $\{N(t), t \geq 0\}$ is a stationary Markov process for which the probabilities $P_{in}(t)$ satisfy the condition

$$\lim_{t \to 0} P_{in}(t) = \delta_{in}, \quad \text{then}$$

(4.1) $$\lim_{t \to \infty} P_{in}(t) = p_{in}$$

exists for all i and n.

A proof of this theorem may be found in Bharucha-Reid [5,p102]. Since for all the queues we have been considering, $N(t)$ is a stationary Markov process, we are assured that the $P_{in}(t)$ approach limits. From the uniform convergence of the generating function $\Pi(z,t)$, defined in equation (2.8), we have

(4.2) $$\lim_{t \to \infty} \Pi(z,t) = P(z)$$

(4.3) where $$P(z) = \sum_{n=0}^{\infty} z^n p_{in} .$$

Since $0 \leq p_{in} \leq 1$ for all i and n, $P(z)$ is convergent when $|z| < 1$.

B. QUEUES WITH BATCH ARRIVALS AND SINGLE DEPARTURES

Applying equation (2.27) to the function $\Pi(z,t)$, we have

$$\lim_{t \to \infty} \Pi(z,t) = \lim_{p \to 0+} p\,\Pi^*(z,p).$$

Using this result, equation (2.13) becomes

$$\lim_{t \to \infty} \Pi(z,t) = \lim_{p \to 0+} \frac{\{h(z)-\mu_0\}\,p P_{i0}^*(p) + \sum_{n=1}^{m-1}\{(1-q)\{h(z)-\mu_0\} + q\sum_{s=n+1}^{m}\mu_s(z^{-s}-z^{-n})\}z^n p P_{in}^*(p)}{f(z) - (\lambda_0+\mu_0+p) + h(z)},$$

Also applying equation (2.27) to the functions $P_{in}(t)$, and using equations (4.1) and (4.2) in the above we obtain

$$(4.4) \qquad P(z) = \frac{\{h(z)-\mu_0\}p_{i0} + \sum_{n=1}^{m-1}[(1-q)\{h(z)-\mu_0\} + q\sum_{s=n+1}^{m}\mu_s(z^{-n}-z^{-s})]z^n p_{in}}{f(z) - (\lambda_0+\mu_0) + h(z)}.$$

For single departures, $\mu_s = \delta_{s1}\mu$, whence

$$(4.5) \qquad P(z) = \frac{\mu(1-z)\,p_{i0}}{z f(z) - (\lambda_0+\mu)z + \mu}.$$

Again applying equation (2.27) to equation (2.21), we find

$$p_{i0} = \lim_{p \to 0+} \frac{p[\xi_1(0)]^i}{\mu[1-\xi_1(0)]}.$$

By theorem 2.1 (p43)

$$(4.6) \qquad p_{io} = \begin{cases} 0 & \text{if } f'(1) \geq \mu, \\ 1 - \frac{f'(1)}{\mu} & \text{otherwise.} \end{cases}$$

Substituting for p_{io} in equation (4.5) we obtain

$$(4.7) \qquad P(z) = \left(1 - \frac{f'(1)}{\mu}\right)\left\{1 - \frac{\{\lambda_0 - f(z)\}z}{\mu(1-z)}\right\}^{-1}, \text{ provided } f'(1) < \mu .$$

Otherwise $P(z) \equiv 0$.

In chapter 2 (p45) we found that $f'(1)$ is the mean arrival rate and μ is the mean departure rate from service. When the inequality **above** is not satisfied, we would expect all the p_{in} to be zero since the queue would almost surely grow without limit. When the inequality $f'(1) < \mu$ is satisfied, we shall say that the queue is "in equilibrium", and we may find the p_{in} by comparing coefficients of z^n in equation (4.7), using equation (4.3). Before doing this for some examples, let us consider another important random variable. Let ν be the time for which a customer waits before commencing service, when the queue is in equilibrium. Then the random variable ν will be called the equilibrium waiting time.

Suppose a customer, A, arrives when there are n (≥ 1) customers in the system. Let $\mu_1, \mu_2, \cdots, \mu_n$ be the service times of the customers before A. Then the equation

$$(4.8) \qquad \text{Prob}(t < \nu \leq t + \delta t) = \text{Prob}(t < \mu_1 + \mu_2 + \cdots + \mu_n \leq t + \delta t)$$

is certainly true if A arrives just after the first customer commences his service period μ_1. That the equation is always true is a consequence of the fact, already mentioned in chapter 1 (p34) that the negative exponential distribution has "no memory".

Since $\mu_1, \mu_2, \cdots, \mu_n$ are identically distributed, independent negative exponential variables, their sum is an E_n variable (see p51). Thus from equation (4.8) we have

$$\text{Prob}\,(t < \nu \le t + \delta t) = \frac{\mu}{\Gamma(n)} (\mu t)^{n-1} e^{-\mu t} \delta t + o(\delta t).$$

Now there may have been any number $n = 1, 2, \cdots$ of customers in the queue when A arrived. Consequently the density function of ν is given by

$$(4.9) \qquad f_\nu(t) = \sum_{n=1}^{\infty} \frac{\mu}{\Gamma(n)} (\mu t)^{n-1} e^{-\mu t} p_{in}.$$

Let $f_\nu^*(p)$ be the Laplace transform of $f_\nu(t)$. Then

$$f_\nu^*(p) = \sum_{n=1}^{\infty} \left(\frac{p+\mu}{\mu}\right)^{-n} p_n = P\left(\frac{\mu}{p+\mu}\right) - p_{io},$$

using equation (4.3). Substituting equation (4.7) for $P(z)$ we find

$$(4.10) \qquad f_\nu^*(p) = \left(1 - \frac{f'(1)}{\mu}\right) \left\{ \frac{\lambda_0 - f\left(\frac{\mu}{p+\mu}\right)}{p - \lambda_0 + f\left(\frac{\mu}{p+\mu}\right)} \right\}.$$

However we have not considered the case when A does not wait.

This will happen if the queue is empty when A arrives, so $F_\nu(t)$, the distribution function of ν, has a jump of magnitude p_{i0}, that is $1 - \frac{f(1)}{\mu}$ at $t = 0$.

Equations (4.7) and (4.10) were obtained, in the case of a general service time distribution by Foster [22] and Wishart [68]. Let us consider the examples which we investigated, for the finite time case, in chapter 2.

1. BATCHES OF CONSTANT SIZE

If $\lambda_r = \delta_{rm}\lambda$, we have the queue with arrivals in batches of constant size m, and $f(z) = \lambda z^m$. The condition for equilibrium is $m\lambda < \mu$, and from equation (4.7)

$$P(z) = \left(1 - \frac{m\lambda}{\mu}\right)\left\{1 - \frac{\lambda z(1-z^m)}{\mu(1-z)}\right\}^{-1}.$$

Comparing coefficients of z^n we get

$$p_{i0} = 1 - \frac{m\lambda}{\mu}$$

$$p_{in} = \frac{\lambda}{\mu}\left(1 - \frac{m\lambda}{\mu}\right)\left(1 + \frac{\lambda}{\mu}\right)^{m-1}, \quad n \leq m,$$

and $$p_{in} = \left(1 - \frac{m\lambda}{\mu}\right)\left\{\sum_{j=0}^{[\frac{n}{m}]} \binom{n-mj}{j}\left(1+\frac{\lambda}{\mu}\right)^{n-mj}\left(\frac{-\lambda}{\lambda+\mu}\right)^j - \sum_{j=0}^{[\frac{n-1}{m}]}\binom{n-1-mj}{j}\left(1-\frac{\lambda}{\mu}\right)^{n-1-mj}\left(\frac{-\lambda}{\lambda+\mu}\right)^j\right\}, \quad \text{if } n > m.$$

The equilibrium waiting time density function has, from equation (4.10), Laplace transform

$$f_\nu^*(p) = \left(1 - \frac{m\lambda}{\mu}\right)\left\{\frac{(p+\mu)^m - \mu^m}{(p-\lambda)(p+\mu)^m + \lambda\mu^m}\right\}$$

<u>when $m = 1$</u>, $\quad f^*(p) = (1 - \frac{\lambda}{\mu}) \frac{\lambda}{p - (\lambda - \mu)}$.

Hence $\quad F_\nu(t) = 1 - \frac{\lambda}{\mu} e^{-(\mu-\lambda)t}$.

The equilibrium queue length probabilities in this case are

$$p_{in} = (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^n.$$

<u>when $m = 2$</u>, $\quad f_\nu^*(p) = (1 - \frac{2\lambda}{\mu})\left\{\frac{p - 2\mu}{p^2 + (\lambda\mu - \lambda)p - \mu(\mu - 2\lambda)}\right\}$

Inverting $f_\nu^*(p)$ by the use of standard tables (see, for example, Erdelyi [17,V1,p229]) we obtain

$$f_\nu(t) = (1 - \frac{2\lambda}{\mu}) e^{-(\mu - \frac{\lambda}{2})t}\left\{\cosh \frac{\sqrt{\lambda(4\mu - \lambda)}}{2} t + \frac{2\mu + \lambda}{\sqrt{\lambda(4\mu - \lambda)}} \sinh \frac{\sqrt{\lambda(4\mu-\lambda)}}{2} t\right\},$$

for the probability density function of ν, and $F_\nu(t)$ has a jump of magnitude $1 - \frac{2\lambda}{\mu}$ at $t = 0$.

The equilibrium queue length probabilities are

$$p_{i0} = 1 - \frac{2\lambda}{\mu},$$
$$p_{i1} = \frac{\lambda}{\mu}(1 - \frac{2\lambda}{\mu})$$
$$\text{and} \quad p_{in} = (1 - \frac{2\lambda}{\mu}) \sum_{r=[\frac{n}{2}]}^{n} (\frac{\lambda}{\mu})^{n-r} \binom{n-r}{r}, \text{ for } n \geq 2.$$

2. BATCHES WITH LOGARITHMIC DISTRIBUTION

If $\lambda_r = \frac{\lambda^r}{r}$, then $f(z) = \log \frac{1}{1 - \lambda z}$. The condition for equilibrium is $\frac{\lambda}{1-\lambda} < \mu$, and the generating function for the

p_{in} is, from equation (4.7),

$$P(z) = \left\{1 - \frac{\lambda}{\mu(1-\lambda)}\right\}\left\{1 - z\left(\log\frac{1}{1-\lambda} - \log\frac{1}{1-\lambda z}\right)/\mu(1-z)\right\}^{-1}.$$

In particular $p_{i0} = \left(1 - \frac{\lambda}{\mu(1-\lambda)}\right)$,

$$p_{i1} = \left(1 - \frac{\lambda}{\mu(1-\lambda)}\right)\frac{1}{\mu}\log\frac{1}{1-\lambda},$$

and $p_{i2} = \left(1 - \frac{\lambda}{\mu(1-\lambda)}\right)\left(\frac{1}{\mu^2}\log^2\frac{1}{1-\lambda} + \frac{1}{\mu}\left(\log\frac{1}{1-\lambda} - \lambda\right)\right).$

Since the p_{in} are complicated for general n , we will not attempt to find an expression for them.

3. BATCHES WITH GEOMETRIC DISTRUBUTION

If $\lambda_r = \lambda^r$, then $f(z) = \frac{\lambda z}{1-\lambda z}$. The condition for equilibrium is $\frac{\lambda}{(1-\lambda)^2} < \mu$. From equation (4.7) we find on simplifying that

$$P(z) = \left(1 - \frac{\lambda}{\mu(1-\lambda)^2}\right)(1-\lambda z)\left\{1 - \lambda\left(1 + \frac{1}{\mu(1-\lambda)}\right)z\right\}^{-1}.$$

Thus $$p_{in} = \frac{\lambda^n}{\mu(1-\lambda)}\left\{1 - \frac{\lambda}{\mu(1-\lambda)^2}\right\}\left\{1 + \frac{1}{\mu(1-\lambda)}\right\}^{n-1}, \quad n \geq 1,$$

with $$p_{i0} = \left(1 - \frac{\lambda}{\mu(1-\lambda)^2}\right).$$

Also from equation (410) it follows that

$$f_2^*(p) = \left\{1 - \frac{\lambda}{\mu(1-\lambda)^2}\right\}\frac{\lambda}{1-\lambda}\left\{p - \left(\frac{\lambda - \mu(1-\lambda)^2}{1-\lambda}\right)\right\}^{-1}.$$

Upon inversion and integration we find that the probability distrubution function of ν is given by

$$F_\nu(t) = 1 - \rho e^{-\mu(1-\lambda)(1-\rho)t},$$

where $\rho = \lambda/\mu(1-\lambda)^2$.

C. STATE-INDEPENDENT PARAMETERS - SINGLE ARRIVALS

If arrivals occur one at a time, then $f(z) = \lambda z$, and so equation (4.4) becomes

$$P(z) = \frac{\{k(z)-\mu_0\}p_{i0} + \sum_{n=1}^{m-1}\left[(1-q)\{k(z)-\mu_0\} + q\sum_{s=n+1}^{m}\mu_s\left(z^{-s}-z^{-n}\right)\right]z^n p_{in}}{\lambda z - (\lambda+\mu_0) + k(z)},$$

which may be written as

$$(4.11) \quad P(z) = \frac{z^m\{k(z)-\mu_0\}\left[p_{i0}+(1-q)\sum_{n=1}^{m-1} z^n p_{in}\right]}{\lambda z^{m+1} - (\lambda+\mu_0)z^m + z^m k(z)} + \frac{q\sum_{n=1}^{m-1}\sum_{s=n+1}^{m}\mu_s\left(z^{m+n-s}-z^m\right)p_{in}}{\lambda z^{m+1} - (\lambda+\mu_0)z^m + z^m k(z)}.$$

Now using theorem 2.3, (see p65) the expression

$$\lambda z^{m+1} - (\lambda+\mu_0)z^m + z^m k(z)$$

has precisely $m-1$ zeros within the region $|z| < 1$ if $\lambda < |k'(1)|$, but m zeros in this region otherwise.

Let us consider first the case when $\lambda < |k'(1)|$. Dividing out the common factor $1-z$ in the numerator and denominator of each part of the right-hand side of equation (4.11) we have

$$(4.12) \quad P(z) = \frac{\sum_{s=1}^{m}\mu_s\sum_{j=m-s}^{m-1} z^j\left[p_{i0}+(1-q)\sum_{n=1}^{m-1} z^n p_{in}\right]}{\lambda(z-\xi_1)(z-\xi_2)\cdots(z-\xi_{m-1})(\eta-z)} + \frac{q\sum_{n=1}^{m-1}\sum_{s=n+1}^{m}\mu_s\sum_{j=m-s+n}^{m-1} z^j p_{in}}{\lambda(z-\xi_1)\cdots(z-\xi_{m-1})(\eta-z)},$$

where $\xi_1, \xi_2, \cdots, \xi_{m-1}$ are the $m-1$ zeros of the denominator within the unit circle, and η is the zero outside this region. The generating function $P(z)$ is convergent within the unit circle. Thus the $\xi_r\ (r=1,2,\cdots,m-1)$ must also be zeros of the numerators in the right-hand side of equation (4.12). Since none of the ξ_r is a zero of the expression $\sum_{s=1}^{m} \mu_s \sum_{j=m-s}^{m-1} z^j$, equation (4.12) thus becomes

$$(4.13) \qquad P(z) = \frac{(1-q)\sum_{s=1}^{m} \mu_s \sum_{j=m-s}^{m-1} z^j p_{i,m-1} + q\sum_{n=1}^{m-1} \mu_{n+1} p_{in}}{\lambda(\eta - z)}.$$

We use the normalising condition

$$(4.14) \qquad \sum_{n=0}^{\infty} p_{in} = P(1) = 1.$$

Putting $z = 1$ in equation (4.13) and using equation (4.14) we find

$$(4.15) \qquad (1-q)\sum_{s=1}^{m} s\mu_s p_{i,m-1} + q\sum_{n=1}^{m-1} \mu_{n+1} p_{in} = \lambda(\eta - 1).$$

In particular, when $q = 0$

$$(4.16) \qquad |k'(1)|\, p_{i,m-1} = \lambda(\eta - 1)$$

Putting equation (4.16) in equation (4.15) we also have

$$(4.17) \qquad \sum_{n=1}^{m-1} \mu_{n+1} p_{in} = \lambda(\eta - 1).$$

Substituting equations (4.16) and (4.17) in equation (4.14) we get

$$(4.18) \qquad P(z) = \frac{\left[(1-q)\sum_{s=1}^{m} \mu_s \sum_{j=m-s}^{m-1} z^j + q|k'(1)|\right](\eta - 1)}{(\eta - z)\,|k'(1)|}$$

Comparing the coefficients of z^n in equation (4.18) and using equation (4.3) we obtain

$$(4.19) \quad p_{in} = \left[\frac{1-q}{|k'(1)|}\sum_{s=0}^{n}(\mu_{m-s}+\cdots+\mu_m)\eta^s + q\right](1-\tfrac{1}{\eta})(\tfrac{1}{\eta})^n.$$

Considering the case when $\lambda \geqslant |k'(1)|$, the denominator of the right-hand side of equation (4.11) has m zeros within the unit circle. However the numerator can have at most $m-1$ zeros in this region, since the numerator of the second part of the right-hand side of equation (4.11) is a polynomial of degree m , which has a zero at $z=1$. Consequently $P(z)$ is not convergent when $\lambda \geqslant |k'(1)|$. In chapter 2 (p66) we saw that λ is the rate at which customers are arriving, and $|k'(1)|$ is the rate at which customers are departing from service. Consequently the condition $\lambda \geqslant |k'(1)|$ expresses the fact that customers are arriving at a rate greater than or equal to the rate with which the server can cope. We would expect the queue to increase without limit in which case $p_{in}=0$ for all finite values of n.

Equation (4.19) is a generalisation of a result obtained by Bailey [3] in the case $q=1$, $\mu_s=\delta_{sm}\mu$. It expresses the equilibrium probabilities in terms of the unique root of the polynomial equation

$$\lambda z^{m+1} - (\lambda+\mu_0)z^m + \sum_{s=1}^{m}\mu_s z^{m-s} = 0$$

outside the unit circle.

The distribution of the equilibrium waiting time for the queue with batch departures will not be considered. Downton [15] obtained this distribution in the special case $q = 1$, $\mu_s = \delta_{sm}\mu$, but it seems that the problem is complicated in general and so we will not pursue it.

D. THE QUEUE WITH STATE-DEPENDENT PARAMETERS

It is a trivial matter to determine the equilibrium queue length probabilities for the non-batch queue with state-dependent parameters. Putting $P_{in}(t) = p_n$ and $\frac{\partial}{\partial t}P_{in}(t) = 0$ in equations (3.1) and (3.2) we have

$$(4.20) \qquad 0 = \lambda_{n-1}p_{n-1} - (\lambda_n + \mu_n)p_n + \mu_{n+1}p_{n+1}, \quad n \geq 1,$$

$$(4.21) \text{ and } \qquad 0 = -\lambda_0 p_0 + \mu_1 p_1 .$$

From equation (4.21) $p_1 = \frac{\lambda_0}{\mu_1}p_0$. Putting this in equation (4.20) and solving recursively for the p_n we deduce

$$(4.22) \qquad p_n = \frac{\lambda_0\lambda_1 \cdots \lambda_{n-1}}{\mu_1\mu_2 \cdots \mu_n}p_0 .$$

We may use the condition $\sum_{n=0}^{\infty} p_n = 1$, and it follows that

$$(4.23) \qquad p_0 = \left\{1 + \sum_{n=1}^{\infty} \frac{\lambda_0\lambda_1 \cdots \lambda_{n-1}}{\mu_1\mu_2 \cdots \mu_n}\right\}^{-1},$$

if the series $\sum_{n=1}^{\infty}\frac{\lambda_0\lambda_1 \cdots \lambda_{n-1}}{\mu_1\mu_2 \cdots \mu_n}$ is convergent. Otherwise $p_n = 0$ for all n.

For the single server queue we have $\mu_n = \mu$, for all n, and the probability density function of the equilibrium waiting time is again given by equation (4.9). We must make a modification for the m-server queue, in that μ_n becomes $m\mu$ if $n \geq m$. We find

$$f_\nu(t) = \sum_{n=m}^{\infty} \frac{\mu m}{\Gamma(n)} (\mu m t)^{n-1} e^{-\mu m t} p_n,$$

(4.24) whence $f_\nu^*(p) = P\left(\frac{\mu m}{p+\mu m}\right) - \sum_{n=0}^{m-1} p_n.$

$F_\nu(t)$ has, in this case, a jump of magnitude $\sum_{n=0}^{m-1} p_n$ at $t=0$.

EXAMPLE - THE SINGLE SERVER QUEUE WITH DISCOURAGEMENT

Let us find the p_n and $f_\nu(t)$ for an example which does not seem to have been treated in the literature. We will assume

$$\lambda_n = \lambda\left(1 - \frac{\alpha}{n+1}\right), \text{ where } -1 < \alpha < \infty,$$

$$\text{and } \mu_n = \mu, \text{ for all } n.$$

This corresponds to the single-server queue in which the probability of an arrival in the interval $(t, t+\delta t]$ decreases as n increases. The situation would arise when prospective customers are discouraged by the sight of a long queue. From equations (4.22) and (4.23) we have

(4.25) $$p_n = \left(1 - \frac{\lambda}{\mu}\right)^{1+\alpha} \left(\frac{\lambda}{\mu}\right)^n \frac{(\alpha)_n}{n!},$$

where $(\alpha)_n = (\alpha+1)(\alpha+2)\cdots(\alpha+n) = \dfrac{\Gamma(\alpha+n+1)}{\Gamma(\alpha+1)}.$

Equation (4.25) is valid provided $\lambda < \mu$, otherwise $p_n = 0$ for all n.

The equilibrium waiting time distribution has, from equation (4.9), probability density function

$$f_\nu(t) = \lambda e^{-\mu t}\left(1-\frac{\lambda}{\mu}\right)^{1+\alpha}\sum_{n=1}^{\infty}\frac{(\lambda t)^{n-1}(\alpha)_n}{(n-1)!\,n!},$$

whence, using the definition of the confluent hypergeometric function (see p53)

(4.26) $$f_\nu(t) = \lambda(1+\alpha)e^{-\mu t}\left(1-\frac{\lambda}{\mu}\right)^{1+\alpha}{}_1F_1(1+\alpha, 2, \lambda t).$$

$F_\nu(t)$ has also a jump of magnitude $\left(1-\frac{\lambda}{\mu}\right)^{1+\alpha}$ at $t=0$.

Another special case arises when $\lambda_n = \frac{\kappa}{n+1}$. However we may incorporate this in the previous example by allowing α to tend to infinity while keeping $\alpha\lambda$ fixed. Let us set $\alpha\lambda = \kappa$. Then equation (4.25) becomes

$$p_n = \left(1-\frac{\kappa}{\mu\alpha}\right)^{1+\alpha}\left(\frac{\kappa}{\mu}\right)^n\frac{(\alpha)_n}{\alpha^n n!}.$$

Now let $\alpha \to \infty$. Then

(4.27) $$p_n = e^{-\frac{\kappa}{\mu}}\left(\frac{\kappa}{\mu}\right)^n\frac{1}{n!}.$$

This formula is valid for all values of κ and μ.

In a similar manner we get from equation (4.26)

$$f_\nu(t) = \kappa e^{-\mu t}e^{-\frac{\kappa}{\mu}}\sum_{n=1}^{\infty}\frac{(\kappa t)^{n-1}}{n!(n-1)!}.$$

Using the definition of the modified Bessel function, that is

$$I_n(x) = \sum_{r=0}^{\infty} \frac{(\frac{1}{2}x)^{2r+n}}{r!\,(n+r)!}$$

(see, for example, Jeffries [30,p574]) we find

$$(4.28) \qquad f_\nu(t) = \sqrt{\frac{k}{t}}\, e^{-\mu t - \frac{k}{\mu}} I_1(2\sqrt{kt}).$$

Upon successive integration of equation (4.28) by parts we find

$$\int_{0+}^{t} f_\nu(u)\,du = 1 - e^{-\frac{k}{\mu}} - e^{-\frac{k}{\mu}-\mu t} \sum_{n=1}^{\infty} \left(\frac{1}{\mu}\sqrt{\frac{k}{t}}\right)^n I_n(2\sqrt{kt}).$$

Since $F_\nu(t)$ has a jump of magnitude $e^{-\frac{k}{\mu}}$ at $t=0$, we finally obtain the expression

$$(4.29) \qquad F_\nu(t) = 1 - e^{-\frac{k}{\mu}-\mu t} \sum_{n=1}^{\infty} \left(\frac{1}{\mu}\sqrt{\frac{k}{t}}\right)^n I_n(2\sqrt{kt})$$

for the probability distribution function of the equilibrium waiting time. This completes our investigation of the asymptotic behaviour of queues.

APPENDIX

<u>THEOREM 1</u> : The Laplace transforms of $\Pi(z,t)$ and $\hat{\Pi}(z,t)$, defined by

$$\Pi^*(z,p) = \int_0^\infty e^{-pt}\Pi(z,t)\,dt \quad \text{and} \quad \hat{\Pi}^*(z,p) = \int_0^\infty e^{-pt}\hat{\Pi}(z,t)\,dt,$$

are each analytic functions of z for fixed p, and p for fixed z, provided the condition

Λ: $|z| \leq 1-\varepsilon$; $Re(p) \geq \delta$, where $0<\varepsilon<1$ and $\delta>0$

is satisfied.

<u>PROOF</u> : Consider the function $z^n P_{in}^*(p)$ for fixed z.

Now $\quad |z^n P_{in}^*(p)| \leq |z|^n \left|\int_0^\infty e^{-pt} P_{in}(t)\,dt\right|,$

$$\leq |z|^n \int_0^\infty e^{-Re(p)t}\,dt,$$

since $0 \leq P_{in}(t) \leq 1$.

Hence $|z^n P_{in}^*(p)| \leq \dfrac{(1-\varepsilon)^n}{Re(p)}$, provided $|z| \leq 1-\varepsilon$.

If $Re(p) \geq \delta$, where $\delta > 0$ is arbitrary, the conditions of Weierstrass's M test (see, for example, Copson [11] p97) are satisfied and so $\sum_{n=0}^{\infty} z^n P_{in}^*(p)$ is uniformly convergent, under condition Λ.

Thus we may interchange the order of summation and integration, and we get

$$\sum_{n=0}^{\infty} z^n P_{in}^*(p) = \int_0^\infty e^{-pt}\Pi(z,t)\,dt$$

Hence, by the analyticity of a uniformly convergent series of analytic functions, $\Pi^*(z,p)$ is analytic in p under condition Λ. Obviously $\Pi^*(z,p)$ is also analytic in z under condition Λ.

The same argument applies to the function $\hat{\Pi}^*(z,p)$.

<u>ROUCHE'S THEOREM</u> : "If $f(z)$ and $g(z)$ are two functions regular within and on a closed contour C, on which $f(z)$ does not vanish and also $|g(z)| < |f(z)|$, then $f(z)$ and $f(z)+g(z)$ have the same number of zeros within C."

(see Copson [11, p119-20], for a proof of this theorem).

<u>THEOREM 2</u> : The function $-f(z)+(\lambda_0+\mu_0+p)-h(z)$ has exactly m zeros inside the contour C: $|z| = 1-\varepsilon$ provided $Re(p) \geq (\lambda_0+\mu_0)\{(1-\varepsilon)^{-m}-1\}$.

<u>PROOF</u> : On C, $z = (1-\varepsilon)e^{i\theta}$, and so

$$|-z^m f(z) - z^m h(z)| = \left|\sum_{r=1}^{\infty} \lambda_r (1-\varepsilon)^{r+m} e^{i(r+m)\theta} + \sum_{s=1}^{m} \mu_s (1-\varepsilon)^{m-s} e^{i(m-s)\theta}\right|,$$

$$< \lambda_0 + \mu_0 .$$

Suppose $\rho = (\lambda_0+\mu_0)\{(1-\varepsilon)^{-m}-1\}$. If $Re(p) \geq \rho$, $\lambda_0+\mu_0 \leq |(1-\varepsilon)^m(\lambda_0+\mu_0+p)|$.
Hence on C when $Re(p) \geq \rho$, $|-z^m f(z) - z^m h(z)| < |z^m(\lambda_0+\mu_0+p)|$.
Now $-z^m[f(z)+h(z)]$ and $z^m(\lambda_0+\mu_0+p)$ are both analytic functions of z within and on C, and the latter does

does not vanish provided $\mu_m \neq 0$. Hence by Rouche's theorem (see above), the function

$$-z^m f(z) + (\lambda_0 + \mu_0 + p) z^m - z^m k(z)$$

has the same number of zeros within C as $(\lambda_0 + \mu_0 + p) z^m$, namely m . If $\mu_n \neq 0$ this function does not vanish at $z = 0$. Hence $-f(z) + (\lambda_0 + \mu_0 + p) - k(z)$ has exactly m zeros inside C, provided $\mathcal{Re}(p) \geq \delta$.

Some results concerning the function $I_n^m(x)$, introduced by Luchak [49].

DEFINITION : (a) $\exp \frac{1}{2} x (z^m + \frac{1}{z}) = \sum_{r=-\infty}^{\infty} z^r I_{-r}^m(x)$,

(b) where $I_n^m(x) = \sum_{r=0}^{\infty} \frac{(\frac{1}{2}x)^{n+(m+1)r}}{(rm+n)!\, r!}$, $m = 1, 2, \cdots$

When $m = 1$, equation (b) becomes $I_n^1(x) = \sum_{r=0}^{\infty} \frac{(\frac{1}{2}x)^{2r+n}}{r!\,(r+n)!}$.

thus $I_n^1(x)$ is the modified Bessel function.

RECURRENCE RELATION : Differentiating equation (a) with respect to z and comparing coefficients of z^r in each side, we have

(c) $$\frac{2r}{x} I_r^m(x) = I_{r-1}^m(x) - m I_{r+m}^m(x).$$

In particular, when $r = 0$

(d) $$I_{-1}^m(x) = m I_m^m(x).$$

DERIVATIVE : Differentiating equation (a) with respect to x and comparing coefficients of z^r in each side, we obtain

(e) $$\frac{\partial}{\partial x} I_r^m(x) = \frac{1}{2}\left[I_{r-1}^m(x) + I_{r+m}^m(x) \right].$$

VALUE AT X=0 : Putting $x=0$ in equation (a) and comparing coefficients of z^r in each side, we get

(f) $$I_r^m(0) = \delta_{r0}.$$

REFERENCES.

1. AITKEN,A.C. (1948): Determinants and Matrices. Oliver and Boyd. 59,63.
2. BAILEY,N.T.J. (1954):A continuous time treatment of a simple queue, using generating functions, J.R.Statist.Soc.,B,v16, 288-291. 30.
3. BAILEY,N.T.J. (1954): On queueing processes with bulk service, J.R. Statist.Soc.,B,v16, 80-87. 26,57,111,121.
4. BATHER,J.A. (1963): Two non-linear birth and death processes, J. Australian Math.Soc., III, 104-116. 75.
5. BHARUCHA-REID,A.T. (1960): Elements of the Theory of Markov Processes and Their Applications. McGraw-Hill. 26,112.
6. BHAT,N.U. (1964): Imbedded Markov chain analysis of single server bulk queues, J.Australian Math.Soc.,lV, 244-263. 36,111.
7. CHAMPERNOWNE,D.G. (1956): An elementary method of the solution of the queueing problem with a single server and constant parameter, J.R.Statist.Soc.,B,v18, 125-128. 30.
8. CLARKE,A.B. (1956): A waiting time process of Markov type, Ann.Math. Stat.,v27, 452-459. 28.
9. CONOLLY,B.W. (1958): A difference equation technique applied to the simple queue, J.R.Statist.Soc.,B,v20, 165-167. 30.
10. CONOLLY,B.W. (1960): Queueing at a single serving point with group arrival, J.R.Statist.Soc.,B,v22, 285-298. 111.
11. COPSON,E.T. (1935): An Introduction to the Theory of Functions of a Complex Variable. Oxford. 47,126,127.
12. COX,D.R. and SMITH,W.L. (1961): Queues. Methuen. 69.
13. DOETSCH,G. (1950): Handbuch der Laplace-Transformation, Birkhauser Verlag, Basel and Stuttgart. 43,85.
14. DOIG,A. (1957): A bibliography on the theory of queues, Biometrika, v44, 490-514. 7.
15. DOWNTON,F. (1955): Waiting time in bulk-service queues, J.R.Statist. Soc.,B,v17, 256-261. 26,57,111,122.
16. ERLANG,A.K. (1909): Probability and telephone calls, Nyt Tidsskr.Mat., B,v20, 33-39. 111.
17. ERDELYI,A. (1954): Tables of Integral Transforms. McGraw-Hill. 55,97,117.

18. ERDELYI,A. (1954):Higher Transcendental Functions. McGraw-Hill. 94.96.97.

19. FELLER,W. (1950): An Introduction to Probability Theory and its Applications. Wiley. 93.

20. FELLER,W. (1940): On the integrodifferential equations of completely discontinuous Markov processes, Trans. Am. Math.Soc.,v48, 488-515. 26.

21. FINCH,P.D. (1963): The single server queueing system with non-recurrent input-process and Erlang service time, J.Australian Math.Soc.,III, 220-236. 111.

22. FOSTER,F.G. (1961): Queues with batch arrivals I,Acta. Math. Hung., Tom. XII, 1-10. 26,41,111,116.

23. FOSTER,F.G. (1953): On stochastic matrices associated with certain queueing processes, Ann. Math.Stat.,v24, 355-360. 75.

24. GAVER,D.P. (1959): Imbedded Markov chain analysis of a waiting-line process in continuous time, Ann.Math.Stat.,v30, 698-720. 36.

25. HAIGHT,F.A. (1960): Queueing with Balking, II, Biometrika,v47, 285-296. 19.

26. HAIGHT,F.A. (1958): Two queues in parallel, Biometrika,v45, 401-410.19.

27. HEATHCOTE,C.R. and MOYAL,J.E. (1959): The random walk (in continuous time) and its application to the theory of queues, Biometrika, v46, 400-411. 29,75.

28. JAISWAL,N.K. (1961): A bulk-service queueing problem with variable capacity, J.R.Statist.Soc.,B,v23, 143-148. 111.

29. JENSEN,A. (1948): An elucidation of A.K.Erlang's statistical works through the theory of stochastic processes, pp23-100 of the memoir of Brockmeyer, Halstrom, and Jensen, The Life and Works of A.K.Erlang. Copenhagen. 111.

30. JEFFRIES, Sir H. and Lady B.(1956): Methods of Mathematical Physics. Cambridge. 53,125.

31. KARLIN,S., and McGREGOR,J. (1957): The differential equations of birth and death processes, and the Stieltjes moment problem, Trans.Am.Math.Soc.,v85, 489-546. 26,75.

32. KARLIN,S., and McGREGOR,J. (1957): The classification of birth and death processes,Trans.Am.Math.Soc.,v86, 366-400. 75,92.

33. KARLIN,S., and McGREGOR,J. (1958): Many server queueing processes with Poisson input and exponential service times, Pacific J. Math.,v8, 87-118. 30,75,101,107.

34. KARLIN,S., and McGREGOR,J. (1958): Linear growth birth and death processes,J.Math.Mech.,v7, 643-662. 75,107.

35. KARLIN,S., and McGREGOR,J. (1962): On a genetics model of Moran,Proc. Cam.Phil.Soc.,v58, 299-311. 75.

36. KARLIN,S., and McGREGOR,J. (1959): Random walks, Illinois J.Maths., v3, 66-81. 107.

37. LINDLEY,D.V. (1952): The theory of queues with a single server, Proc. Camb.Phil.Soc.,v48, 277-289. 111.

38. KEILSON,J. (1962): Non-stationary Markov walks on the lattice, J. Math.andPhys.,vXL1, 205-211. 69,75,101.

39. KEILSON,J. (1963): On the asymptotic behaviour of queues, J.R.Statist. Soc.,B,v25, 464-476. 111.

40. KEILSON,J. (1962): The use of Green's functions in the study of bounded random walks with application to queueing theory, J.Math. and Phys.,vXL1, 42-52. 28.

41. KEILSON,J. (1964): Some comments on single server queueing methods and some new results,Proc.Camb.Phil.Soc.,v60, 237-252. 36,57,73.

42. KENDALL,D.G. (1951): Some problems in the theory of queues, J.R. Statist.Soc.,B,v13, 151-185. 30,74,111.

43. KENDALL,D.G. (1949): Stochastic processes and population growth, J.R.Statist.Soc.,B,v11, 230-264. 93.

44. KENDALL,D.G. (1953): Stochastic processes in the theory of queues,
45. Ann.Math.Stat.,v24, 338-354. 101,111.

46. KHINCHINE,A.JA. (1933): Uber die mittlere Dauer des Stillstandes von Maschinen, Mat.Sbornil,v40, 119-123. 111.

47. KIEFER,J., and WOLFOWITZ,J. (1955): On the theory of queues with many servers, Trans.Am.Math.Soc.,v78, 1-18. 101.

48. LEDERMANN,W., and REUTER,G.E. (1954): Spectral theory for the differential equations of simple birth and death processes,Phil. Trans.Roy.Soc., London,A,v246, 321-369. 26,28,30,51,74.

49. LUCHAK,G. (1956): The solution of the single-channel queueing equation characterised by a time-dependent Poisson-distributed arrival rate and a general class of holding times, operations Res.,v4, 711-732. 41,49,128.

50. LUCHAK,G. (1958): The continuous time solution of the equations of the single channel queue with a general class of service-time distributions by the method of generating function, J.R.Statist. Soc.,B,v20, 176-181. 26,36,41,49,50.

51. LUKE,Y.L. (1962): Integrals of Bessel Functions. McGraw-Hill. 51,57.

52. MEIXNER,J. (1934): Orthogonale Polynomsysteme mit einer besondere Gestalt der Erzengendem Funktion, J.London Math.Soc.,v9, 6-13.95.

53. MILLER,R.G. (1959): A contribution to the theory of bulk queues, J. R.Statist.Soc.,B,v21, 320-337. 111.

54. MORAN,P.A.P. (1963): Some general results on random walks, with genetic applications, J.Australian Math.Soc.,vIII, 468-479. 29,75.

55. MORSE,P.M. (1958): Queues, Inventories and Maintenance. Wiley. 29,30,111.

56. NIELSON,K.L. (1962): Differential Equations. Barnes and Noble. 98.

57. PARZEN,E. (1962): Stochastic Processes. Holden-Day. 16.

58. POLLACZEK,F. (1930): Uber eine Aufgabe der Wahrscheinlichkeitstheorie, part II,Math.Z.,v32, 729-750. 42,111.

59. SAATY,T.L. (1961): Elements of Queueing Theory. McGraw-Hill. 29,101

60. SACK,R.A. (1963): Treatment of the non-equilibrium theory of simple queues by means of cumulative probabilities, J.R.Statist.Soc.,B, v25, 457-463. 75.

61. SHOHAT,J.A., and TAMARKIN,J.D. (1943): The Problem of Moments. Mathematical Surveys,Volume 1. Waverly. 80,85.

62. TAKACS,L. (1962): Introduction to the theory of Queues. Oxford. 29,42,111.

63. TODHUNTER,I. (1882): The Theory of Equations. Macmillan. 59.

64. VON SYDOW,L. (1958): Some Aspects on the variations in traffic intensity, Teleteknik, 58-64. 28.

65. WHITTAKER,E.T., and WATSON,G.N. (1920): A Course of Modern Analysis. Cambridge. 47.

66. WIDDER,D.V. (1946): The Laplace Transform. Princeton. 103.

67. WISHART,D.M.G. (1960): Queueing systems in which the discipline is "last-come, first served", Operations Res.v8, 591-599. 19.

68. WISHART,D.M.G. (1956): A queueing system with chi-squared service-time distribution, Ann.Math.Stat.v27, 768-779. 26,116.

69. SZEGO,G. (1939): Orthogonal Polynomials, Amer.Math.Soc.Coll.Publcns, v23. 88.